Decision weights for experimental asset prices based on visual salience

Devdeepta Bose,[1] Henning Cordes,[2] Sven Nolte,[3]
Judith C. Schneider,[2] Colin F. Camerer,[1*]

February 1, 2020

Abstract

Using a machine-learning algorithm that can predict visually salient portions of images, we construct decision weights based on visually salient parts of a price chart. We analyze these weights in three experimental studies that vary in the realism of the price path images and task complexity that investors are exposed to. We find that these decision weights are strongly predictive of future investments. We conclude that visual salience captures attention paid to historical returns. Visual salience goes beyond overweighting returns at the tails of the historical distribution as in the models of Barberis et al. (2016) and Bordalo et al. (2012). Moreover, we find that visual salience affects investment decisions independently from recency effects.

JEL classification: G4, G12
Keywords: Decision Weights, Eye-Tracking, Machine-Learning Algorithm, Visual Salience

1 Introduction and Motivation

Benartzi and Thaler (1995) in their seminal work on the equity premium puzzle propose that investors form expectations about future returns of an investment by using the historical distribution of annual stock market returns. Newer models suggest applying prospect theory to

*We thank Nicholas Barberis, Lawrence Jin, Alex Imas, Elise Payzan-LeNestour, and seminar and conference participants at Warwick Business School, Radboud University, University of Münster, Caltech, Southwestern Economic and Behavioral Economics 2019, Experimental Science Association 2019, Mid-Adlantic Meeting on Behavioral and Experimental Economics 2019, and the Boulder Summer Conference on Consumer Financial Decision Making 2019 for helpful comments and suggestions.
[1]California Institute of Technology, Corresponding author is devbose@caltech.edu
[2]University of Münster, Universitätsstr. 14-16., 48143 Münster, Germany
[3]Institute for Management Research, Radboud University, Nijmegen, The Netherlands

the historical return distribution (Barberis et al., 2016) or weighting historical returns by their salience compared to a benchmark return (Bordalo et al., 2012; Cosemans and Frehen, 2017). Both these types of models assign decision weights to individual returns from the past to form an expectation of future returns.

These classes of decision weights have strong psychologically-driven foundations. We add to this literature by relying on psychophysiologically motivated measures of attention for the decision weights. Attention is a scarce resource for investors during financial decision-making. Events that are likely to be attention-getting are thought to affect what investments people make (Barber and Odean, 2007; Hillert et al., 2014; Mohrschladt and Schneider, 2018). In decision weights as in Barberis et al. (2016) or Bordalo et al. (2012), the role of attention is accounted for by overweighting returns at the tails of the historical distribution. It is natural to conclude that investors pay more attention to these extreme returns. However, we contend that these models do not take into account the *visual* salience of price charts capturing attention effects towards specific returns beyond overweighting the tails.

The most common representation of a stock's performance is a visualization of historical prices via a chart. We argue that the decision weights investors assign to individual past returns are driven by the allocated attention. We propose that initial attention towards specific returns is determined by the visual salience of respective points on the visualized price path. An effective way to measure visual attention is by recording visual fixations and eye movements (saccades). Early studies using measures of visual fixation in economics showed that what people looked at, and what they did not, could explain deviations of strategizing from equilibrium predictions (Camerer et al., 1993).

More recent research has extended this insight using a neuroscientifically-inspired algorithm (Cornia et al., 2018, Salience Affective Model (SAM)) to predict which points of an image are more visually salient than others. SAM is trained on actual eye-tracking data. Since price

charts are an ubiquitous source of information for investors, we claim that SAM can identify returns in these price chart images that are more (or less) visually salient, and assign higher (or lower) decision weights to these returns. Investor's valuations of historical returns can then be modeled by using a preference functional consisting of the product of the decision weights and a value function. We propose that visual salience enters the functional by affecting the decision weights. This approach enables us to compare our visual salience model with the existing models of Barberis et al. (2016) and Bordalo et al. (2012).

We collect data from three experimental studies to evaluate the performance of our visual salience model. The three studies vary with respect to the degree of realism of the depicted price charts and the degree of control over the predictions of the different models. The purpose of Study I is to explore if our visual salience model has predictive power for investments based on empirical price charts that investors are exposed to in reality. We find that our model is in contrast to the models of Barberis et al. (2016) and Bordalo et al. (2012) able to significantly predict invested amounts even without specifying a particular value function.

Study II is designed to isolate the impact of the effect of temporal ordering of returns on predictability. Barberis et al. (2016) accomodate for a recency effect by adding an exponential decay function to CPT which accounts for overweighting more recent returns (see also Bordalo et al., 2019). Our data suggest that there are at least two coexisting, independent effects: a recency effect, and a visual attention and salience effect. Depending on the context, investors might either rely more on recent returns or on the most visually salient returns.

Study III considers the simplest possible depiction of price charts controlling for a myriad of factors which could influence investment decisions like image complexity or task complexity. Moreover, in this study we have control over the environment and attention of the participants. We show that even in this simple and controlled setup, our model is able to predict future investments.

1.1 Related Literature

This paper adds to four different strands of literature: First, we contribute to the literature that explores how the framing or presentation of historical returns can stimulate different biases and heuristics within investors, and influence their decision making. Recent experimental evidence suggests that price path shapes influence asset attractiveness and investment behavior (Grosshans and Zeisberger, 2018; Nolte and Schneider, 2018). Moreover, the representation of the price chart, i.e., the visual properties (e.g., scaling) of charts influence risk perception and return expectations (Huber and Huber, 2018). Most information brochures display a visualization of an asset's price path. The average investor is easily overwhelmed by the complexity and variety of offered products. Thus, an easily accessible graphical presentation is expected to have a significant impact on decision making. There are several experimental studies which address the impact of price path presentation on investment behavior. Glaser et al. (2019) analyze how representation of returns vs. prices shape individual forecasts of asset performance. Weber et al. (2005) find that the presentation format of historical asset returns influences expected asset volatility. Diacon and Hasseldine (2007) find that the presentation format of asset returns has a significant effect on subjects' perceptions of risk and return. Glaser et al. (2007) associate beliefs in mean reversion or trend chasing with framing effects. Asset forecasts vary depending on whether investors are presented with returns or with prices, and the authors attribute these differences to the representativeness heuristic. They conclude that surveys on asset forecasts should rely on price presentation rather than returns, see also Czaczkes and Ganzach (1996). Grosshans and Zeisberger (2018) present subjects with variously shaped price paths, and ask them to record their levels of satisfaction if they had hypothetically held such an asset. They conclude that the shape of the historical asset price path might shape future investment behavior through reference point effects. Our visual salience model shows that visual salience of certain returns in a historical price path generate attention towards those returns, which is what drives

the predictability of invested amounts in all three experimental studies.

Second, this paper adds to the literature applying visual salience to economic decision making. Outputs from visual salience algorithms have been shown to predict economic choice in other domains. Towal et al. (2013) find that objective visual properties such as saliency, and subjective properties such as value, are both interactively useful in predicting decisions in multiple-alternative forced-choice environments, e.g., choosing snacks from a vending machine. Li and Camerer (2019) find that visual salience of focal points in images are useful in predicting winners in coordination games and hider-seeker games. Thus, visual salience has been used in fields such as consumer choice and empirical game theory. Our paper extends the scope of its use by introducing this concept to the field of behavioral finance.

Third, this paper adds to the literature which proposes different theoretical models on how to integrate past returns into forming expectations about future returns. A centerpiece of any investment choice is to form expectations about the future performance of investment opportunities at the time of judgment. These expectations crucially depend on available information and how this information is processed. Barberis et al. (2001, 2016) argue that prospect theory values based on past returns predict future returns. One key assumption in their model is that the presentation format of the price charts is irrelevant. The ordering of returns is only captured in an extension of their model, by an exponential overweighting of the most recent returns, i.e., only temporal ordering of the returns is relevant. In another class of models by Bordalo et al. (2012) and applied in Cosemans and Frehen (2017) and Gödker and Lukas (2017), the magnitude of the return difference to some reference value determines the weight assigned to the return. Thus, this version of salience is context dependent, i.e., dependent on the other available assets. These salient (extreme) returns attract attention and predict future returns. One limitation of these models is that the salience rank of a past return is independent of the actual probability of its occurrence. Moreover, the temporal ordering of returns and presentation

format are irrelevant in these models. In contrast, our proposed model based on the Salience Attentive Model (Cornia et al., 2018) assumes that the weighting depends only on visual features. Thus, presentation format and temporal ordering of returns both matter. The returns close to salience points are overweighted. While in the models of Barberis et al. (2016) and Bordalo et al. (2012), the weighting depends on specific functional forms, the weighting implied by visual salience is parameter free and depends on feature extraction by a machine learning algorithm.

Lastly, our paper adds to the literature on the role of attention in financial decision making. Given the almost infinite amount of stock market information readily available to any investor, paying attention to every piece would exceed individuals' cognitive abilities (Kahneman, 1979). Only a few stocks end up in the choice set of investors. A firm which catches more investor attention is more likely to be included in the investors' portfolio (Odean, 1999; Busse and Green, 2002; Barber and Odean, 2007). A stock can attract attention through media coverage (e.g., Fang and Peress, 2009), lottery-like features (e.g., Kumar, 2009), high daily trading volume (e.g., Barber and Odean, 2007), extreme daily returns (e.g., Bali et al., 2011; Ungeheuer, 2017), daily newpapers winners and losers (e.g., Kumar et al., 2019), or related events (e.g., Koester et al., 2016, for earnings surprises). All these papers study indirect measures of attention. An exception is Da et al. (2011) who provide evidence that stock-related Google search volume reflects the attention paid by private (not sophisticated) investors as they gather information using Google. Thus, Google search volume indices provide a timely measure of firm-level, private investor attention. Which features of these stocks determine their attractiveness also depends on which feature grabs most attention. Inattention to information affects choices of private investors as well as analysts (Koester et al., 2016). While these approaches are various their common idea is to determine the choice set of stocks investors are paying attention to. In contrast, within this choice set, we look at attention towards path characteristics of an asset measured by using eye-tracking, i.e., by using a specific computational model of where attention

is predicted to be allocated to.

In summary, we propose a model that estimates salience based on actual measurements (eye-tracking) while previous models on salience miss the visual impact of return representation.

2 Theoretical Framework

Let X be a random variable with realizations $x_1, ..., x_m$ where $m \in M$ denotes either past realizations $k \in K$, or future realizations $s \in S$. M can be interpreted as the state space which describes the choice problem under risk. Since the entire state space M and possible future realizations $s \in S$ are unknown to the decision maker (DM), she relies on past realizations to help evaluate what is possible in the future. Barberis et al. (1998) argue that such an extrapolation is psychologically rooted in the "representativeness heuristic", a concept introduced by Kahneman and Tversky (1972, 1973). This heuristic leads the DM to think that the limited sample of past realizations is representative of the full distribution of possible outcomes, weighted by the likelihood that the DM associates with each past realization. In our setting, investors use a combination of past returns and decision weights to evaluate the appeal of investing in the future. The investors evaluate X according to:

$$V(X) = \sum_{k \in K} \pi_k v(x_k), \tag{1}$$

where $v(\cdot)$ is a continuous, strictly increasing value function, and π_k represent the decision weights associated with historical realizations that occurred with objective probability p_k. In the following sections, we will explore how visual salience affects the decision weights that the investors assign to the past realizations, and how our model compares against the decision weights used in the canonical studies of Barberis et al. (2016) and Bordalo et al. (2012). We also explore how our results are impacted when value functions are replaced by correlations between returns and decision weights.

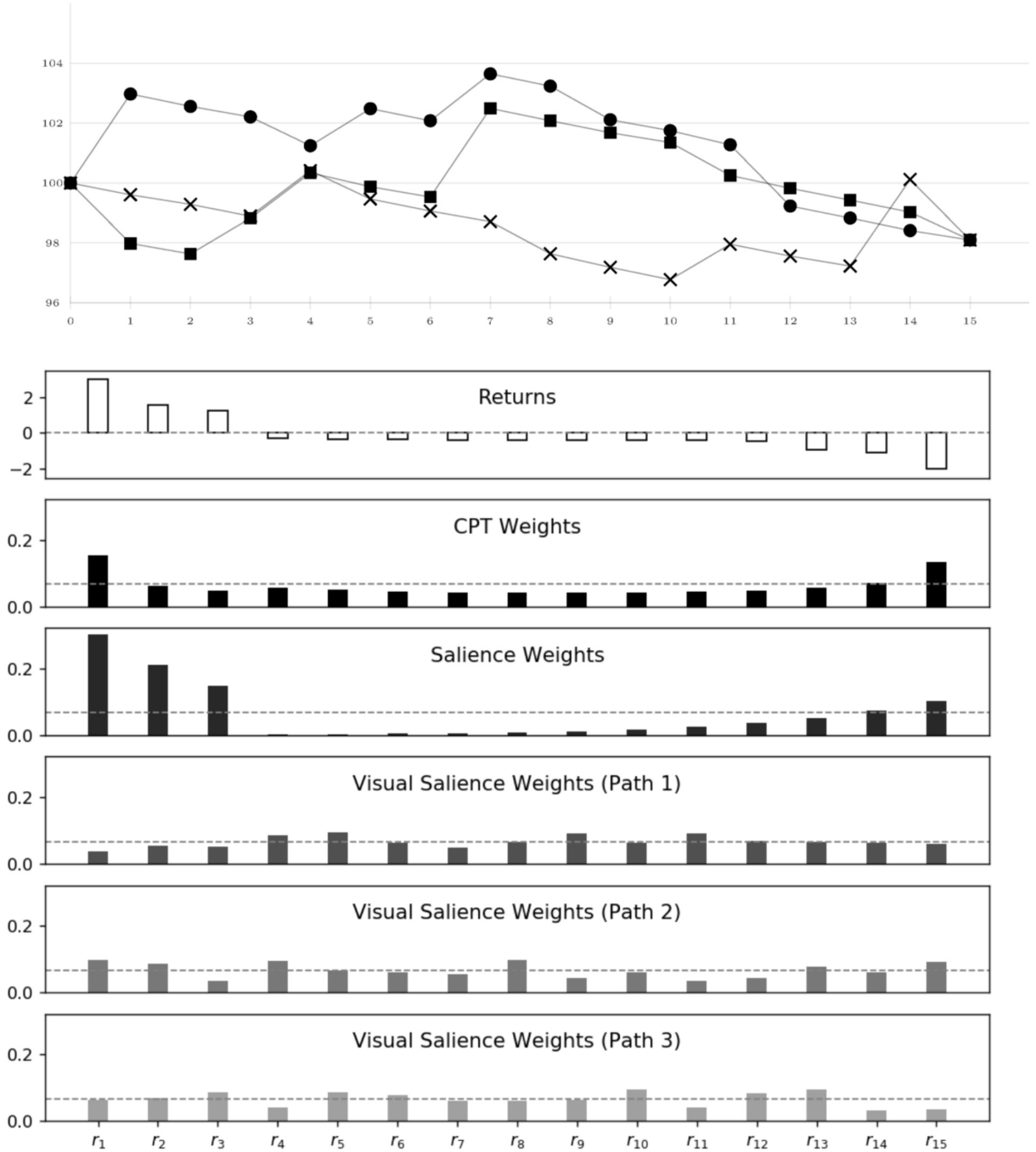
2.1 Decision Weights

The general purpose of decision weights is to formalize how investors feel about probabilities (Wakker, 2010). In an applied decision problem like an investment decision, it is common to employ a two step procedure to determine the decision weights. First, a mathematical functional that transforms objective probabilities into decision weights is developed. This functional is motivated by psychological insights and usually based on a set of parameters. These parameters are calibrated to observed behavior in a second step. This procedure underlies the decision weights in prospect theory (Tversky and Kahneman, 1992) or salience theory (Bordalo et al., 2012).

Our visual salience decision weights do not follow this standard procedure. Instead, objective probabilities are directly distorted by *measured* attention weights. The only functional form we impose is that the visual salience decision weights (attention weights $\times$ objective probabilities) have to add up to one. No further specific parametrization is needed. Recent research in neuroeconomics motivates the use of attention weights by showing that subjects in the laboratory attach more weights to attributes that they attend to more when making binary or multi-attribute choices, as well as choices among risky lotteries (Krajbich et al., 2010; Krajbich and Rangel, 2011; Mormann and Frydman, 2017; Mullett and Stewart, 2016).

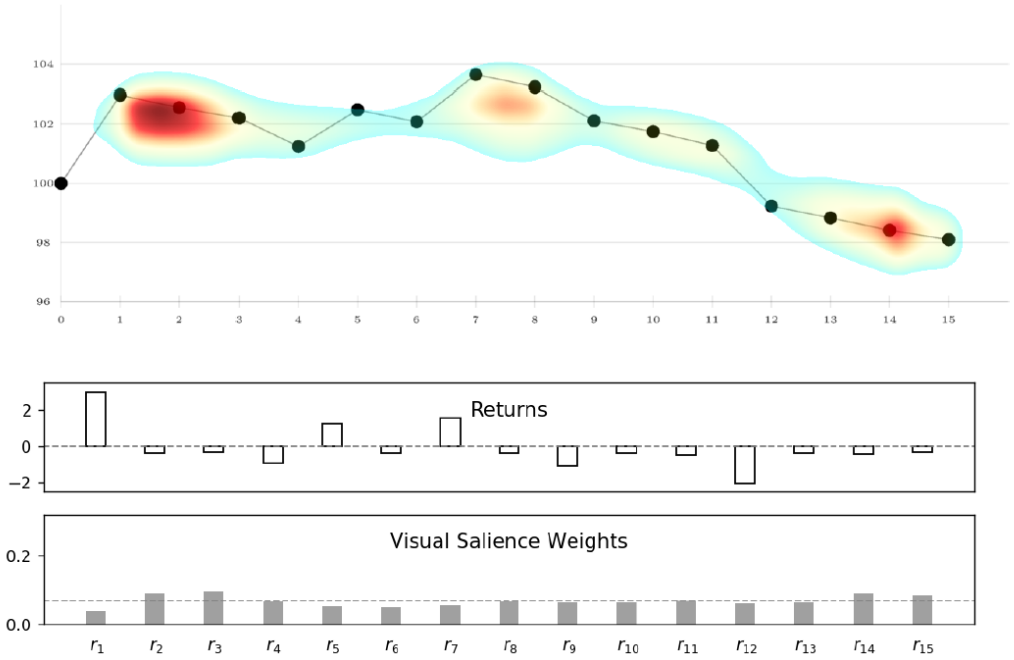
To illustrate how CPT, salience theory, and our visual salience approach determine decision weights, consider the following example. Let x_t be the price for every period following a stationary AR(1) process with $x_{t+1} = \rho x_t + \varepsilon_t$ with $\rho = 0.6$. Three exemplary associated price paths for 15 time periods of historical realizations are depicted in the upper panel of Figure 1. All three paths are constructed from the identical 15 returns, only the temporal ordering of the returns differs. As a result, the visual appearance and the shapes of the price paths also differ. Associated returns are depicted in descending order below the price charts. Empirical frequencies for each of the 15 returns are $p_k = 1/15$, $k \in \{1, ..., 15\}$ and will be treated as

Figure 1: CPT and Salience



The figure shows three price paths of 15 periods historical realizations constructed from the returns in the graph below. The returns are the result of stationary process $x_{t+1} = 0.6x_t + \varepsilon_t$. The two graphs below present decision weights based on CPT and salience theory. The bottom three paths are visual salience decision weights for path 1 (dots), 2 (crosses), and 3 (squares).

Figure 2: CPT and Salience



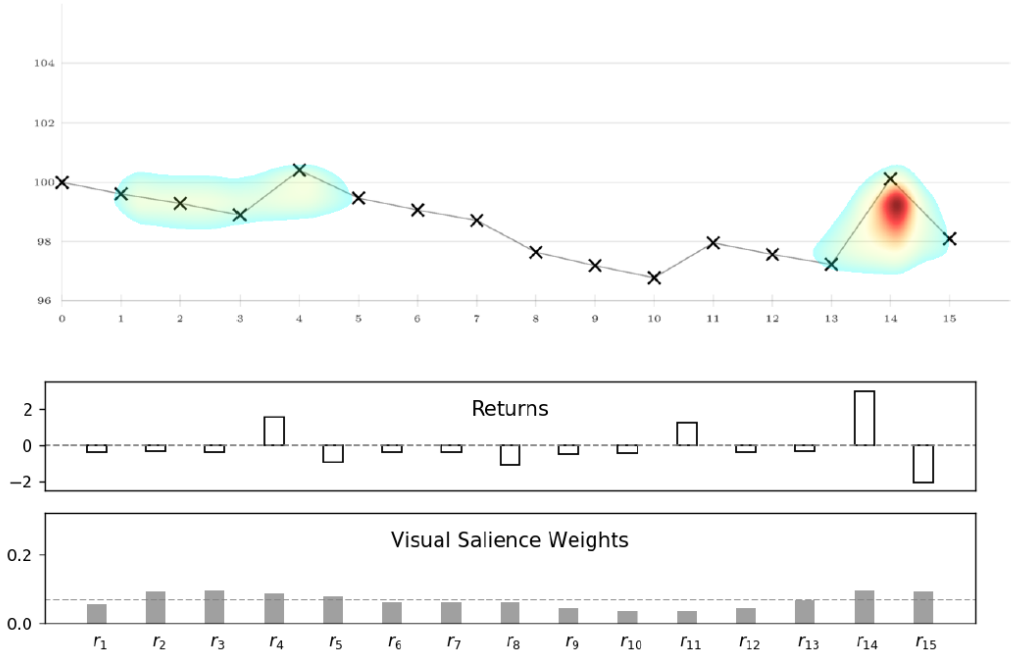
The figure shows the first price path from Figure 1 with an overlay of the heat map returned by SAM. The darker the overlay the more attention is allocated to that area of the price path. The returns in the order of occurrence and the resulting visual salience weights are depicted below.

objective probabilities.

Probability weighting as in CPT transforms the corresponding objective probabilities using the following functional. The historical returns are separated into gains and losses relative to a pre-specified reference point (e.g., the status quo), and sorted in ascending order. The weighting function for each return is applied to the cumulative probability for all the returns that are smaller than or equal to the reference point (for losses), or all returns that are larger than or equal to the reference point (for gains).[1]

[1]We assume that $\pi^+(0) = \pi^-(0) = 0$ and $\pi^+(1) = \pi^-(1) = 1$.

Figure 3: CPT and Salience



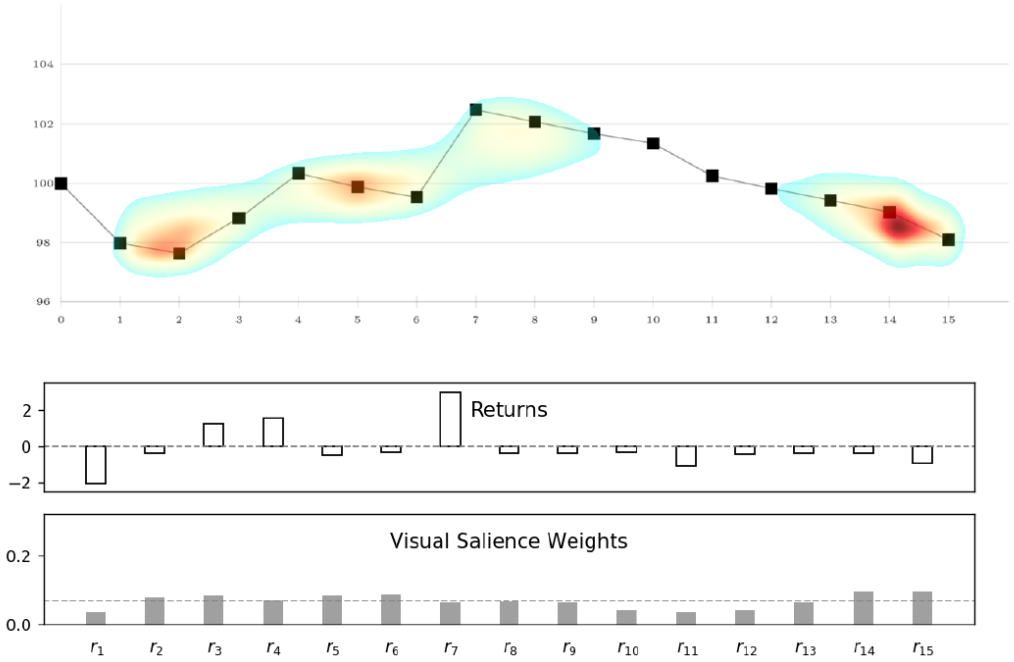
The figure shows the second price path from Figure 1 with an overlay of the heat map returned by SAM. The darker the overlay the more attention is allocated to that area of the price path. The returns in the order of occurrence and the resulting visual salience weights are depicted below.

$$\pi_k^{\text{CPT}} = \begin{cases} w^+(p_k + \ldots + p_n) - w^+(p_{k+1} + \ldots + p_n), \ 0 \leq k \leq n \text{ and} \\ w^-(p_{-m} + \ldots + p_k) - w^-(p_{-m} + \ldots + p_{k-1}), \ -m \leq i \leq -1 \end{cases} \tag{2}$$

$$\text{and } w^+(p) = \frac{p^{\delta_+}}{(p^{\delta_+} + (1-p)^{\delta_+})^{\frac{1}{\delta_+}}} \text{ and } w^-(p) = \frac{p^{\delta_-}}{(p^{\delta_-} + (1-p)^{\delta_-})^{\frac{1}{\delta_-}}} \quad \delta(0,1] \tag{3}$$

Based on experimental evidence, probability weighting functions are predominantly inverse S-shaped, which means concave for small probabilities and convex for moderate and high probabilities (Abdellaoui, 2000). The original parametrization for this functional is $\delta^+ = 0.69, \delta^- = 0.61$ (Tversky and Kahneman, 1992), but the optimal parametrization depends on the underlying decision problem and elicitation method used (Wu and Gonzalez, 1996). Moreover, several alternative specifications of the probability weighting functional itself are discussed in the lit-

Figure 4: CPT and Salience



The figure shows the third price path from Figure 1 with an overlay of the heat map returned by SAM. The darker the overlay the more attention is allocated to that area of the price path. The returns in the order of occurrence and the resulting visual salience weights are depicted below.

erature (Goldstein and Einhorn, 1987; Prelec, 1998; Diecidue et al., 2009). While the exact decision weights change depending on the parametrization or weighting function, the general idea remains: Extreme returns are overweighted. This can be seen from the third graph in Figure 1. Moderate returns are underweighted while extreme returns are overweighted. One characteristic that is important to note is that CPT decision weights do not depend on the temporal order of returns.

In salience theory (Bordalo et al., 2012), decision weights are derived from salience weights. Salience overweights returns that have a large magnitude of difference from some pre-specified reference value, and underweights returns that are close to this reference value. This idea of salience is operationalized by assigning each return a salience rank, depending on its difference

from the reference value.

$$\pi_k^{Sal} = p_k \times h_k \text{ with the salience weight } h_k = \frac{\nu^{\kappa_k}}{\sum_i \nu^{\kappa_i} p_i} \quad \nu \in (0, 1] \tag{4}$$

where κ_k denotes the salience rank of x_k, which is measured by:

$$\kappa_k = \sigma(x_k, \bar{x}_k) = \frac{|x_k - \bar{x}_k|}{|x_k| + |\bar{x}_k| + \theta} \text{ where } \bar{x}_k \text{ is a reference value}$$

These decision weights rely on the idea that investors attend to differences rather than absolute values (Kahneman, 2003). They have the properties of *ordering* and *diminishing sensitivity* where salience increases in the distance from the reference value, and salience decreases as the reference value gets farther from zero. Often the reference value is the aggregated stock market return or a similar benchmark (Cosemans and Frehen, 2017). For a price path, one possible reference value is the average of all historical returns. This implies that returns that are dramatically different from this mean are salient in the minds of investors, and are thus overweighted when aggregating historical returns into future expectations. For our analyses, we use the parameter values suggested in Bordalo et al. (2012) of $\nu = 0.7, \theta = 0.1$.[2] An illustration can be found in the lower graph of Figure 1. The reference value is close to the moderate realized returns that is why they receive the lowest weight. The reference value makes salience theory context-dependent, and thus flexible and applicable to many economic decisions. At the same time, eliciting the reference value is difficult, and there exists no standard procedure, so exact salience weights are somewhat arbitrary. Also, just like CPT weights, they do not depend on the temporal order of returns.[3]

[2]The smaller the ν, the more weight is put on the extreme returns (farther from the reference value).

[3]For both CPT and salience weights, we present in Section 1, Figure 1 of the Online Appendix analyses where we vary the parameters, the CPT weighting function, and the reference value for salience. The impact on the decision weights is as expected and rather small.

For decision weights based on visual salience, we rely on direct measures of (visual) attention via tracking of investors' eye movements. To make this approach feasible, we leverage recent technological advancements in the computer vision literature and use a neural network model called the Salience Attentive Model (SAM). This model uses direct measures of attention through eye-tracking fixations to train an algorithm that looks at pixels within the training images that correlate with a high number of eye fixations, and then attempts to find similar features in out-of-sample images. For our application, we input images of price charts into the algorithm, and SAM outputs a greyscale image assigning higher values (close to 1) to pixels that contain features that are more visually salient and lower values (close to 0) to pixels that contain features that are less visually salient. We aggregate the visual salience values for all the pixels associated with each price, and average the values across consecutive prices to obtain a decision weight for each return.

Visual salience probability weighting is defined as:

$$\pi_k^{\text{VS}} = p_k \times l_k, \sum_k \pi_k = 1, \text{ with the visual salience weight } l_k = \frac{\gamma_k}{\sum_i \gamma_i p_i} \tag{5}$$

$$\text{and } \gamma_k = \frac{SAM(P_{k-1}) + SAM(P_k)}{2}$$

where $SAM(P_k)$ denotes the visual salience of price P_k as predicted by SAM.

Figures 2 - 4 illustrate the decision weights and the outcome of SAM for the three price charts used in Figure 1. Obviously, decision weights are now dependent on the order of return realizations. The decision weights assign higher probabilities to areas of the price paths where SAM predicts higher attention. Thus, these decision weights do not depend on the extremity of the return but on the attention it receives, i.e., it is not necessarily the highest return which receives the highest weight but is dependent on the occurrence and attention allocation. While

in Figure 2 the highest return gets assigned a small weight, in Figure 3 the highest return gets assigned a relatively high weight.

We can compare decision weights obtained by the three theories by calculating distance and correlation measures. We employ three metrics for this purpose: the sum of absolute distances from empirical frequencies ($\sum_{t=0}^{T} |\pi_k - p_k|$), the sum of absolute distances between decision weights ($\sum_{t=0}^{T} |\pi_k^i - \pi_k^j|$), and the linear correlation between decision weights. Each of these metrics is calculated on a path level. Table 1 provides the three metrics for the price paths used in this section's stylized example. Decision weights derived from salience theory are based on the salience rank of returns, so for any given number of return periods T, the decision weights and therefore distances from empirical frequencies associated with a particular rank will be identical. Thus, there is no variation in distances across different paths of the same T. The same holds true for CPT decision weights, with the slight difference that because they are cumulated separately for gains and losses, differences in the number of positive and negative returns across paths will lead to marginally different absolute differences from the empirical frequency across paths with identical T. Correlations between salience theory and CPT decision weights will also be substantial because they both assign larger decision weights to more extreme returns. In contrast, visual salience theory allows for any distribution of decision weights and is independent from return rankings or T. Thus, we see a wider dispersion of decision weights across paths and a low correlation with the other two theories. Section 10 of the Online Appendix provides distributions and summary statistics for the three metrics and the price paths used in Study I. Properties are consistent with those from our toy example, i.e., our visual salience decision weights have a wider distribution and are uncorrelated to the other theories' decision weights. Correlations range evenly from slightly negative to slightly positive, enabling a good separation between visual salience and other theories across paths.

To prove that visual salience decision weights indeed carry relevant information, let us con-

Table 1: Distance and correlation metrics

Sum of absolute distance from empirical frequencies

	Path1	Path2	Path3
CPT	0.36	0.36	0.36
Sal	1.01	1.01	1.01
VS	0.20	0.29	0.26

Sum of absolute distance between theories

	Path1	Path2	Path3
CPT - Sal	0.78	0.78	0.78
CPT - VS	0.49	0.38	0.53
VS - Sal	1.14	0.93	1.09

Correlations between theories

	Path1	Path2	Path3
CPT - Sal	0.68	0.68	0.68
CPT - VS	-0.47	0.48	-0.39
VS - Sal	-0.67	0.33	-0.04

Distance and correlation metrics for our three sample paths and the theories by Barberis et al. (2016) (CPT), Bordalo et al. (2012) (Sal), as well as our visual salience theory (VS). The table on the left displays the sum of absolute distances from empirical frequencies ($\frac{1}{15}$), the center table shows the sum of absolute distances between the different theories' decision weights, and the table on the right contains the degree of linear correlation between the respective decision weights.

sider an example based on the data collected in Nolte and Schneider (2018). In this study investors are primed with four price path shapes with the same returns but different temporal ordering, and then asked for their willingness to invest. In Figure 2 of the Online Appendix (Section 2), we depict the two price paths with the highest and lowest visual salience value. The two price paths are designed such that the return realizations, i.e., the expected value is the same for both. If visual salience weights indeed have forecasting ability, the price path which results in the highest (lowest) expected value based on salience decision weights (left graph, right respectively) should also show the highest (lowest) invested amount. Indeed, for the two paths the average invested amount is equal to 65.48% (left graph) and to 54.76% (right graph). This difference of more than 10 percentage points is both economically and statistically significant ($p < 0.001$).

2.2 Value Functions and Correlation Measure

To setup the full model as outlined in Equation 1, we still need to define the value function. We rely on the CPT value function which incorporates three different aspects that are psychologically motivated: reference dependence, loss aversion, and diminishing sensitivity. The functional form we use is:

$$v(x_s) = \begin{cases} x_s^{\alpha} & \text{for } x_s \geq 0 \\ -\lambda(-x_s)^{\alpha} & \text{for } x_s < 0 \end{cases} \tag{6}$$

with $\lambda > 1$ and $\alpha < 1$.

All x_s are expressed as gains or losses relative to a reference point. Thus, the DM evaluates relative returns rather than the absolute value of returns when making her decision. How reference points are formed remains an open question. For now, we encode positive returns as gains and negative returns as losses.[4]

Loss aversion is the idea that people are more sensitive to losses than equally sized gains. This is captured in the CPT value function by λ, the loss aversion parameter where $\lambda > 1$ implies that losses loom larger than gains.

Finally, the parameter α measures the diminishing marginal sensitivity that the DM has towards large losses or large gains in the value function. As the basic parameter setup, we will stick to the common choice of $\alpha = 0.88$ and $\lambda = 2.25$.

Both the decision weights and the value functions affect the evaluation of the historical price path. Since our visual salience measure only enters the functional through the decision weights, we propose an additional measure based on correlations that serves as a proxy for the direct influence of decision weights on utility independent of the value function.[5] Note that we can rewrite Equation 1 as:

$$V(X) \propto E(\pi_k v(x_k))$$

$$V(X) \propto E(\pi_k)E(v(x_k)) + Cov(v(x_k), \pi_k)$$

[4]For price paths Baucells et al. (2011) and Nolte and Schneider (2018) have proposed that the reference point is some weighted combination of the initial, final, average, highest, and lowest price.

[5]When the number of decision weights is much larger than the number of possible returns, we artificially depress any covariational/correlational structure. In Study III, we explore risky assets with only two possible returns. In this case, to have a meaningful setup weights should collapse to the number of outcomes. This only holds for the model of Bordalo et al. (2012), but not for the specification of CPT proposed by Barberis et al. (2016) or for our visual salience weights. Thus, we will not use the correlational measures for the analyses in Study III.

Under salience weighting and visual salience weighting, the first term is proportional to $\sum_{k\in K}\pi_k \sum_{k\in K} v(x_k)$ as we do not need to separate the decision weights by gains and losses. In both specifications, the decision weights add up to 1 which implies that $E(\pi_k) \propto \sum_{k\in K}\pi_k = 1$. Thus, the first term directly reflects the influence of the value function on utility. For CPT, this only holds for situations in which only gain or loss states exist.

To capture the direct influence of different decision weights on $V(X)$, we pay attention directly to the covariance term. Based on an infinite taylor series expansion of $v(\cdot)$ around some constant, we can use the correlational measure $Corr(x_k, \pi_k)$ as a proxy. For details of the derivation we refer to Section 3 of the Online Appendix.

3 Saliency Attentive Model (SAM)

Our hypothesis is that when investors look at a price chart, their gaze is drawn towards salient regions through attentive mechanisms. There is a long literature from neuroscientists and computer scientists attempting to emulate this process by using saliency prediction models to predict gaze locations. We use one such model created by the AImage Lab at the University of Modena and Reggio Emilia called the Salience Attentive Model (SAM) (Cornia et al. (2018)). SAM is a saliency prediction neural network model consisting of three main blocks: a feature extraction Convoluted Neural Network (CNN) that identifies features of input images that coincide with a high incidence of eye fixations, a Long Short Term Memory (LSTM) attention based feature encoding network that refines and balances the weights assigned to different high and low level feature maps when making predictions for out-of-sample images, and a prior learning network which accounts for the tendency of humans to fixate on the center region of an image.

3.1 Off-the-shelf SAM

SAM has been trained and validated on a combination of four image datasets: SALICON (Jiang et al., 2015), MIT300 (Judd et al., 2012), MIT1003 (Judd et al., 2009), and CAT2000 (Borji and Itti, 2015). Together, these datasets include more than 23,000 images that highlight complex everyday scenes containing common objects in their natural context.[6] Sources for these images include Microsoft COCO, Flickr, and LabelMe. The images in these datasets consist of natural indoor and outdoor scenes and were originally used to train machine learning algorithms for object identification. SAM is trained and validated against the eye-tracking fixations of human subjects on these images to predict visually salient regions in out-of-sample images, such as the price charts we base our experiments on.

The baseline SAM trained on images of natural indoor and outdoor scenes identifies features of the images that correlate highly with a higher number of fixations made by human observers. The identified features could be low-level (bottom-up) such as horizontal and vertical lines, or brighter colors, or shifts in contrast and shapes, but could also be high-level (top-down) like faces, trees or animals etc.[7] SAM identifies 512 features from the images in a training set, and continually refines these features using an attentional LSTM to identify an optimized set of features that are correlated with eye fixations (or lack thereof). The trained SAM then attempts to identify these refined features in out-of-sample images in the test set to predict which pixels within an image are more likely to be looked at, and which pixels are likely to be ignored. SAM makes predictions about portions of an image that will be fixated on within the first few seconds of interaction. We hypothesize that these initial seconds are enough for agents to form

[6]Some categories used to separate the images in these datasets include: Action, Affective, Art, Black & White, Cartoon, Fractal, Indoor, Inverted, Jumbled, Line Drawing, Low Resolution, Noisy, Object, Outdoor Man-made, Outdoor Natural, Pattern, Random, Satellite, Sketch, and Social.

[7]It should be noted that some features identified by the CNN in SAM are a matrix of numbers that do not have any obvious interpretation in the English language, which makes the underlying algorithm somewhat "black-box" in nature.

a heuristic about potential future returns when shown images of historical price charts. This is very similar to the ideas outlined in Barberis et al. (2016).

The SAM output after processing an input of a price chart is a corresponding greyscale image of the same resolution as the input image, with brighter patches (values closer to one) showing visually salient pixels, and darker patches (values closer to zero) showing visually non-salient pixels. We then identify the pixels within each input price chart, and add up the visual salience numbers of the corresponding pixels in the output image. These visual salience numbers are then normalized into decision weights so that they add up to one (see Section 2.1).

3.2 SAM and Price Paths

The off-the-shelf SAM is trained on images that are domain-neutral and should capture general human traits during image perception. That is why we expect that the decision weights should be predictive even in a domain such as financial price paths that is substantially dissimilar from the original objectives of SAM. However, we wanted to explicitly test this assertion by comparing SAM's predictions to actual human subject fixations for a sample of price path images. To operationalize this, we conducted an experiment with 57 participants at the California Institute of Technology. We tracked participants' eye-movements and, more importantly, their fixations while they were presented with price paths on a computer monitor. Following the original training procedure for MIT300 images (Judd et al., 2012), participants were provided with the following instructions, "You will see a series of 60 price path images. Look closely at each price path image. After every three seconds, a new price path image will automatically appear. After viewing the price path images you will have a memory test: you will be asked to identify whether or not you have seen particular images before."

There was a memory test afterwards to motivate users to pay attention. Images were only shown for three seconds each to capture the parts to which initial attention is allocated. Par-

ticipants were presented with a total of 60 price path images. 30 of these paths were selected to match those in Study III to allow for a more thorough analysis of price path images with obvious up-down and down-up trends. The other 30 paths were drawn from actual empirical price paths with distinct visual properties like the ones in Study I.

The participants were tested individually using an EyeLink 1000 Plus (SR Research, Osgoode, ON, Canada) eye-tracker. Monocular eye movements were recorded at 500Hz and fixations were identified by the eye tracker using velocity algorithms. The experiment was displayed on a widescreen Dell monitor (1280 x 1024 resolution). Participants were seated approximately 28 inches away from the monitor to allow for accurate eyetracking and comfortable gameplay. All subjects underwent a 13 point calibration and validation cycle of the eyetracker before proceeding with the experiment trials. Stimulus presentation was controlled by MATLAB using Psychtoolbox extensions (Brainard, 1997; Pelli, 1997).

We obtain a continuous fixation map for each price path image from the eye tracking data by convolving a Gaussian filter across fixation locations of all observers. A Gaussian filter is used to spread out each fixation to the pixels that are within 1 degree of visual angle (or approximately 35 pixels) to match with the area that an observer sees at high focus around the point of fixation. Figure 5 gives two examples of the groundtruth density maps from our experiment on the left, and the SAM-predicted heatmaps overlaid with the price path participants looked at on the right. By visual inspection, the figures confirm that predictions and fixations look very similar. Moreover, we use these groundtruth density maps along with the binary fixation data as inputs for different metrics to evaluate quantitatively how well SAM predicts salience points within price paths.

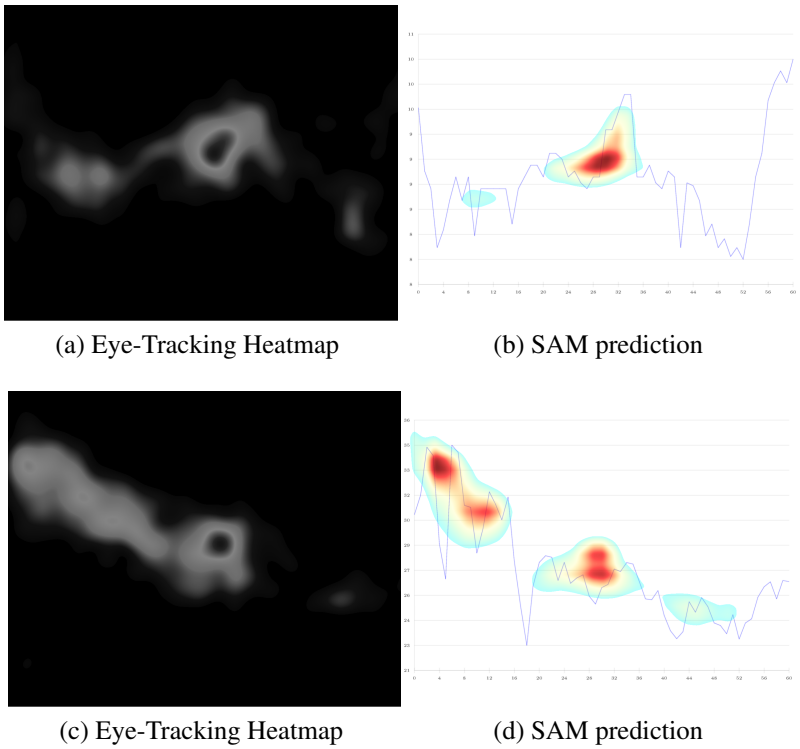
For an overview of suitable metrics, we refer to Le Meur and Baccino (2013) and Bylinskii et al. (2018). We use six different metrics. Three belong to the group of metrics based on receiver operating characteristics (ROC). In signal detection theory, the ROC measures the

trade-off between true and false positives at various discrimination thresholds. The area under the ROC curve (AUC) gives us the degree of similarity of two salience maps. A value of one indicate a perfect classification while a value of 0.5 indicates classification as by any random picture. We also include the normalized scanpath saliency (NSS) which uses the SAM saliency maps and the set of fixations as inputs. The idea is to measure the saliency values at fixation locations along subjects' scanpaths. Le Meur and Baccino (2013) define 0.3 as a good correspondence between a model-predicted saliency map and the observer's scanpath. The simplest measure we include is the pearson correlation coefficient (CC). Lastly, as a measure of dissimilarity we take KL-divergence which is a measure from information theory capturing the overall dissimilarity between two probability distributions. Table 2 presents the results. We compute metrics based on the participants' fixations maps for price path images and the salience maps predicted by SAM. As a benchmark, we compare this to SAM's performance on the set of domain-neutral images on which it was originally trained and validated. Finally, we compare participants' salience maps with heatmaps where (for each price path image) fixations were randomly assigned. While there is some drop-off in SAM's performance when applied to the domain of price path images, it is not significant. Also, SAM performs significantly better than random assignment of fixations when making predictions. Thus we conclude that SAM, although not trained on price path images, captures well what people perceive as salient within a price path.

4 Experiments

We collect experimental data to evaluate the performance of the visual salience model. The studies vary with respect to the degree of realism of the depicted price charts and the degree of control over the predictions of the different models. Two studies (Study I and II) are settled on an experiment run on Amazon Mechanical Turk (MTurk). The third study (Study III) is based

Figure 5: Heatmaps



(a) Eye-Tracking Heatmap (b) SAM prediction

(c) Eye-Tracking Heatmap (d) SAM prediction

The figure shows two examples of the groundtruth density maps from our experiment on the left, and the SAM-predicted heatmaps where we also show the underlying price paths participants looked at on the right.

on a lab experiment conducted at a German university. The purpose of Study I is to explore if our visual salience model has predictive power for investments based on empirical price charts that investors are exposed to in reality. Study II is designed to isolate the impact of the effect of temporal ordering of returns on predictability. Study III considers the simplest possible depiction of price charts controlling for a myriad of factors which could influence investment decisions like image complexity, task complexity etc. We also provide full access to the data generating process behind the price paths such that a rational choice could be made without too much effort. Moreover, in this study we have perfect control over the environment and attention of the participants. In all three studies, the start price is fixed at 100 monetary units to keep the scaling of the abscissa constant and to limit psychological framing effects (e.g., Glaser et al.,

Table 2: Evaluation metrics

	AUC-J.	AUC-B.	sAUC	CC	NSS	KL
SAM (domain-neutral)	0.87	0.78	0.70	0.78	2.34	1.27
SAM vs fixations (price paths)	0.81	0.73	0.70	0.52	1.17	1.58
Random vs fixations (price paths)	0.50	0.50	0.50	0.07	0.002	12.57
Range (worst-best)	0-1	0-1	0-1	0-1	0-∞	0-∞

The table reports six evaluation metrics. We compare SAM's performance on the domain-neutral images it was trained on against SAM's performance against human eye-fixations for price path images. We also use random assignment of fixations as a comparison to see how well SAM is performing for price path images. We include three metrics using receiver operating characteristics analysis and calculating the area under the curve: AUC-J. (Judd et al., 2009), AUC-B.(Borji et al., 2012), and sAUC (Borji et al., 2013). We also include the Pearson Correlation Coefficient (CC), the Normalized Scanpath Saliency metric (NSS) as in Peters et al. (2005), and Kullback-Leibler divergence (KL).

2007; Huber and Huber, 2018).

4.1 MTurk Experiment

For Study I and II, we recruited 500 participants from MTurk which is an online platform that allows researchers to post small tasks that require a human to perform. Potential participants can browse through the lists of small tasks and can self-select into the tasks of their interest. The large available population on MTurk ensures heterogeneity in participant characteristics and a speedy data collection for an acceptable amount of money. Our task offers the participants a fixed and a variable completion fee, the latter depending on their choices within the experiment. The fixed fee was set to 2 US-$ and the variable part was on average 0.94 US-$. Participants took a mean of 12 minutes to complete the experiment, while 95% of all completion times were in the interval between 5 and 30 minutes. After completing the experiment, each participant received an individual, alpha-numerical code to verify their participation. The participant typed in the code on the MTurk Webpage and after having verified their participation, we authorized the payment.

The experiment consists of three parts: an introduction with a small tutorial to ensure that participants understand the general task, the task itself and a final questionnaire.

We show each participant ten different pre-selected price charts. Eight price charts stem from a set based on stocks from the Center for Research in Security Prices (CRSP) in 2017 presenting the input data for Study I. Two charts are based on constructed price charts simulated via a Geometric Brownian Motion (GBM), presenting the input data for Study II. For each price chart, participants decide how attractive the price chart is, what return they expect over the next 12 months, how risky they deem an investment into this stock, and finally and most importantly, how much they want to invest over the next year from their initial endowment of 1000 monetary units into this stock vs a safe bank account.[8] We emphasize that the chart can give the participant an idea about the risk and return of investing over the next year but that the return over the next year can also be much higher or lower. Participants make their investment decisions by using a slider which allows them to invest every amount between zero and 100% of their initial endowment. To avoid anchoring effects, there is no pre-selected slider position at the beginning of each period. Instead, participants need to click at any position of the slider to activate it; afterwards, they can move the slider position freely to decide on their investment. The variable payment is based on the investment task, i.e., we incentivize this task separately.[9] The participant learns that the computer will randomly select one of his 10 investment decisions after having completed the final questionnaire. The participants final wealth is the value of the selected stock investment after one year plus the money in the bank account. The return of the selected stock investment after one year is based on this stock's actual return in the subsequent year, i.e., CRSP data for 2018 for Study I and simulated data for Study II. The participant received 0.001 US-$ for every monetary unit of his final wealth of the selected decision in the

[8]For a detailed description of the questions and an example of the interface, we refer to Section 4 of the Online Appendix.

[9]This is one of the provisions we include to account for the recent MTurk data quality debate (Kennedy et al., 2018).

investment task.

For Study I, we preselect 1000 price charts based on all 8,453 companies inferred from CRSP for 2017-2018. We first remove double entries, firms with incomplete price data, firms with negative prices, penny stocks (i.e., prices below 5 US-$), and lastly incomplete stock/year combinations. In total, that leaves us with 4,246 stocks. We then form ten times ten portfolios based on two criteria. The first criteria is the stock's return within 2017. The second criteria is the degree of convexity and concavity of the price path. Paths that are classified as highly convex (concave) by our applied score tend to exhibit specific visual features, most prominently troughs (peaks) followed by rather long winning (losing) streaks. Previous experiments indicate that historical price paths with similar characteristics attract the attention of investors (Grosshans and Zeisberger, 2018; Nolte and Schneider, 2018). For a detailed description of the applied score we refer to Online Appendix Section 5. First each stock is allocated to one decile based on its return in 2017. Second, within each decile, every stock is assigned to a convexity score decile. We then randomly pick 10 stocks from each bucket. We ensure that each participant chooses on eight different stock charts where the order is randomized. In total, every price chart is evaluated four times ensuring that each participant evaluates each chart no more than once.

For Study II, we select 300 constructed price charts. We start constructing our price paths by simulating 252 daily returns based on a GBM. We implement 3 different parameter constellations for the GBM where the base case implies a mean of 6% and a volatility of 20%, then we decrease the mean in one specification to -6% and fix realized skewness of the drawn path to zero in the third specification.[10] More precisely, we draw one possible path realization for each parameter constellation from the underlying lognormal distribution. The drawn path is the base for constructing 100 different price paths shapes: We split the price path into twelve sub-

[10]In the experiment we include two additional parameter specifications with an extreme high volatility and an extreme low volatility. To leave one parameter fixed and increase the homogeneity in the price paths, we exclude these in this analysis.

periods of equal length mirroring twelve months and rearrange the subperiods randomly. Thus, all shapes of price paths are identical with respect to daily returns and returns of the twelve sub-periods. Moreover, they are also identical for any sub-periods that are multiples of twelve. This implies that individuals having CPT-preferences as described in the basic model of Barberis et al. (2016), or Salient-preferences as in Bordalo et al. (2012) or Cosemans and Frehen (2017) would assign the same value to every price path.

Besides demographic questions, we include questions on risk taking, and general traits of decision making. Moreover, we control for attention and understanding by including a set of quiz questions in the aftermath of the experiment. For details on the questionnaire, we refer to the Online Appendix Section 3.

As Online Appendix Section 8.1 Table 1 shows, the recruited MTurk sample matches the US population exhibiting a somewhat overrepresentation of high-education groups and younger individuals. This is consistent with previous literature documenting that MTurkers are actually quite representative of the population of U.S. internet users (Ross et al., 2010; DellaVigna and Pope, 2017).

4.1.1 Results: Study I

Study I addresses how our visual salience model performs compared to the CPT-framework and the salience-framework in predicting invested amounts of our participants relying on empirical price charts. Table 3 presents the regression results of valuations on invested amounts for all 1000 CRSP-based price charts included in the MTurk study. We approximate the value for all three models by the correlation of the historical return and the different decision weights as outlined in Section 2. This isolates the effect of different value functions for the different models from the implications of the decision weights. Model (1) of Table 3 only includes the correlation measure for the visual salience decision weight. As controls we include in each regression the

Table 3: Regressions for IA, Study I: Correlation Measure

	(1) IA [%]	(2) IA [%]	(3) IA [%]	(4) IA [%]	(5) IA[%]
Corr(Return, VS weights)	0.670***			0.635***	0.706***
	(0.235)			(0.237)	(0.249)
Corr(Return, CPT weights)		0.289		0.984**	1.147**
		(0.242)		(0.434)	(0.534)
Corr(Return, Salience weights)			-0.0271	-0.249**	-0.310**
			(0.0689)	(0.121)	(0.146)
Return	0.297***	0.294***	0.295***	0.291***	0.367***
	(0.0206)	(0.0203)	(0.0204)	(0.0206)	(0.0160)
Stdv	-0.130***	-0.137***	-0.131***	-0.140***	-0.184***
	(0.0152)	(0.0158)	(0.0152)	(0.0164)	(0.0159)
Skewness	0.00859	-0.0183	0.0159	-0.0303	-0.0340
	(0.0112)	(0.0257)	(0.0193)	(0.0266)	(0.0349)
Constant	0.00617***	-0.0144	-0.00142	-0.0562**	-0.117**
	(0.00216)	(0.0121)	(0.00362)	(0.0275)	(0.0488)
Observations	4000	4000	4000	4000	4000
R^2	0.162	0.160	0.160	0.163	

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

The table reports invested amounts on the correlation between current return, visual salience decision weight, CPT decision weight, and the salience decision weight. Control variables are the return of the price path (Return), its standard deviation (Stdv), and its skewness (Skewness). In Model (5) we use a Tobit regression to account for accumulation of participants who did not invest in that stock. We use fixed effects to control for individual specific effects. All variables are standardized for easy comparison of coefficient sizes and to enhance readability of the table. Robust standard errors clustered on participant level are reported in parentheses.

return, standard deviation and skewness of the different price charts to account for heterogeneity in stock characteristics. The results confirm that there is a significant and positive relation between our measure and the invested amount. For CPT weights, this relation is insignificant and positive (Model (2)) while for salience weights the relation is insignificant and negative (Model (3)). Model (4) presents the results of a regression where we include all three measures. Note that the correlation between CPT weights and salience weights is above 90%, thus we face multicollinearity. Despite this, the coefficient of our measure remains positive and significant. Model (5) is a Tobit regression to account for some clustering of invested amounts at 0 which has almost no impact on our visual salience coefficient. In our MTurk sample, we see that in 13% of all decisions nothing was invested and 0.2% of the participants never invested. In our Online Appendix Section 8, we run further regression analyses to account for MTurk specific issues. One regression excludes participants who fail attention and comprehension filters, and another regression uses the attractiveness question as a control to check the consistency of the answers. The results stay qualitatively the same. It is also noteworthy that in line with economic intuition, the return has a positive and significant coefficient while the standard deviation has a negative and significant coefficient for all models.

As additional analyses, we present in Table 4 the same five regression as in Table 3 but for a CPT value function. Table 4 shows that with a CPT value function, our visual salience model is the only model which positively and significantly predicts invested amounts. For CPT and salience decision weights, all coefficients remain insignificant changing between positive and negative sign across regressions.

To sum up, regardless of the specification of the value function, our model is the only one which consistently predicts invested amounts significantly and positively. Our results thus provide evidence for the idea that when investors look at price charts, their gaze is drawn towards salient regions through attentive mechanisms. We believe that we add to the understanding of

Table 4: Regressions for IA, Study I: CPT Value Function

	(1) IA [%]	(2) IA [%]	(3) IA [%]	(4) IA [%]	(5) IA [%]
Visual Salience	0.0955***			0.106**	0.109**
	(0.0357)			(0.0415)	(0.0474)
CPT		0.0184		-0.0489	-0.0588
		(0.0590)		(0.0809)	(0.0955)
Salience			-0.0117	0.00768	0.00821
			(0.0242)	(0.0285)	(0.0346)
Return	0.252***	0.289***	0.297***	0.264***	0.342***
	(0.0255)	(0.0296)	(0.0206)	(0.0315)	(0.0327)
Stdv	-0.0453	-0.115**	-0.136***	-0.0776	-0.126*
	(0.0364)	(0.0567)	(0.0180)	(0.0599)	(0.0723)
Skewness	0.0139	0.00406	0.0187	0.0235	0.0241
	(0.0111)	(0.0211)	(0.0209)	(0.0247)	(0.0280)
Constant	-4.60e-09***	-4.68e-09***	-4.66e-09***	-4.55e-09***	-0.0501
	(2.31e-11)	(5.55e-11)	(1.96e-11)	(9.01e-11)	(0.0356)
Observations	4000	4000	4000	4000	4000
R^2	0.162	0.160	0.160	0.162	

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

The table reports invested amounts on the Visual Salience model, the CPT model (Barberis et al., 2016) and the Salience model (Bordalo et al., 2012) where we assume gain-loss utility with curvature $\alpha = 0.88$, a reference point ρ of zero, and a loss aversion coefficient of $\lambda = 2$. Control variables are the return of the price path (Return), its standard deviation (Stdv), and its skewness (Skewness). In Model (5) we use a Tobit regression to account for accumulation of participants who did not invest in that stock. We use fixed effects to control for individual specific effects. All variables are standardized for easy comparison of coefficient sizes and to enhance readability of the table. Robust standard errors clustered on participant level are reported in parentheses.

what drives attention by focusing on the *visual* component. While we do not claim that this is the only source of what drives decision weights, we think it is should be included in models of decision making when expectations are formed taking the past realizations as a proxy.

4.1.2 Results Study II

In Study II, we aim at disentangling the effect of visual salience from pure temporal ordering like a recency effect, i.e., recent historical returns are more likely to be representative of returns in the immediate future than returns that are from a long time ago (e.g., Bordalo et al., 2019). Recency effects capture a fundamental psychological factor that approximates the speed of memory decay, and has some empirical support in extrapolation based models (e.g., Cassella and Gulen, 2018). While CPT-values and salience-values are meaningless in Study II, Barberis et al. (2016) suggest a modification of the original CPT to accomodate recency effects. They introduce a recency parameter ρ which underweights (geometrically) returns that are further in the past.

More specifically, t_k is the number of months ago that return r_k was realized, the multiplicative term in front of the value function is ρ^{t_k} normalized by the weighted sum of all ρ. In the robustness test of Barberis et al. (2016), ρ equals $0.9, 0.85, 0.8$. Barberis et al. (2016) also argue that investor's quick, passive reaction to a chart is likely to be based on the chart as an integral whole, with the early part of the chart affecting the investor just as much as the later part. Our decision weights basically capture the weight investors assign to returns which are visually salient, while ρ captures a recency effect in the weighting of the returns. To see whether a modified CPT-value has predictive power for our constructed price charts, we run additional regressions. Table 5 presents the results where in Models (2)-(4) we include $\rho = 0.95, 0.85, 0.5$. Model (1) includes our visual salience weights with a CPT-value function. Interestingly, only for a very strong recency effect does the CPT-value become significant. Including both CPT-

Table 5: Regressions for IA, Study II: Recency Effect

	(1) IA[%]	(2) IA [%]	(3) IA [%]	(4) IA [%]	(5) IA [%]
Visual Salience	0.205*				0.217**
	(0.107)				(0.107)
CPT (0.95)		0.00417			
		(0.0418)			
CPT (0.85)			0.0362		
			(0.0393)		
CPT (0.50)				0.0961*	0.104*
				(0.0583)	(0.0585)
6, 20	0.0139	0.0556	0.0496	0.0706	0.0252
	(0.113)	(0.116)	(0.111)	(0.105)	(0.109)
6, 20, skew	0.0985	0.130	0.113	0.128	0.0909
	(0.103)	(0.116)	(0.112)	(0.0997)	(0.102)
Constant	-0.0710	-0.0627	-0.0521	-0.0473	-0.0530
	(0.0618)	(0.0713)	(0.0701)	(0.0665)	(0.0648)
Observations	600	600	600	600	600
R^2	0.030	0.011	0.015	0.024	0.045

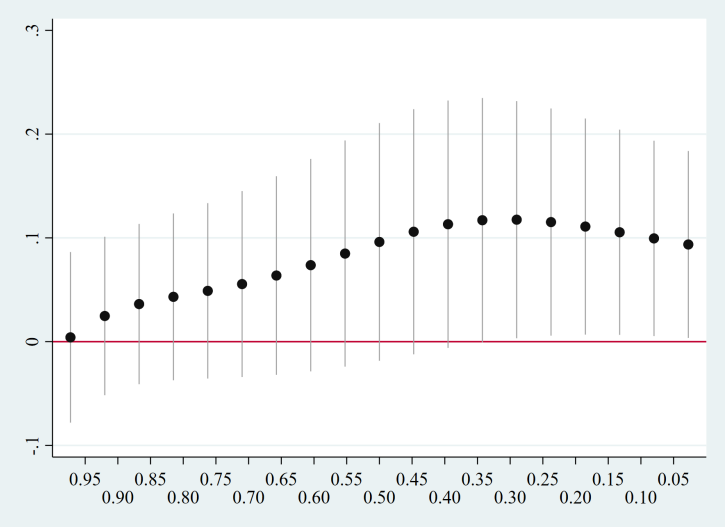
Standard errors in parentheses

* p<0.1, ** p<0.05, *** p<0.01

The table reports invested amounts on the Visual Salience Model with CPT value function (Model(1)) and a CPT model with recency parameter $\rho = 0.95, 0.85, 0.50$, Models (2)-(4). In Model (5) we run the regression with the Visual Salience value and the CPT value ($\rho = 0.5$) as explanatory variables. We also include controls for the parameter constellation of the constructed price charts, with $\mu = -6\%$ and $\sigma = 20\%$ as reference category.

and our visual salience-value, shows that neither model is redundant. To determine the exact parameter value ρ for which the CPT value becomes significant, we run regressions varying ρ from 0.95 to 0.05. Figure 6 depicts the results by illustrating the coefficient and confidence intervals for regressions of CPT with varying recency parameter on invested amounts. For very low parameter values, the coefficient becomes insignificant again. Moreover, our model remains significant also when including the CPT model with recency weighting. This gives rise to the question whether in fact investors exhibit a recency and a visual attention/salience bias. Empirical research has shown that the 52-week high has predictive power for subsequent returns (George and Hwang, 2004). Translated to our model this could mirror a salience weighted event, while the general tendency to overweight most recent events is an independent effect. A further investigation of this relation is left for future research. But our results clearly demonstrate that to fully understand investment choices, a combination of different approaches could be fruitful.

Figure 6: CPT with Recency weighting



The figure shows the coefficient and confidence intervals for regressions of CPT with varying recency parameter on invested amounts.

4.2 Laboratory Study

4.2.1 Experimental Setup

In Study III, we run a laboratory experiment to analyze how well our visual salience model fares in a controlled environment with highly simplified investment options. The goals of this study are threefold: First, we reduce task complexity to a minimum by making the available asset a binary lottery which yields a positive or negative return with known outcomes and probabilities. This implies that participants face no uncertainty regarding the underlying processes that determine future asset prices and returns. Moreover, the investment task and information provided are comparably easy to process. Consequently, any biases or distortions of the investment decision caused by the complexity of the task are also minimized. Second, we make the visualization process easy to understand. Participants receive graphical information about the lottery outcomes in form of a price chart. The price chart is a depiction of previous periods' realizations of the binary lottery, and participants see how the price path updates after each realization. This makes the fact that a price path is merely a visualization of lottery outcomes (asset returns) apparent to participants. Given that the participants attended advanced statistic classes during their studies, they should be able to fully understand the mechanisms behind how a price path develops over time. Third, we run the experiment with a homogeneous student sample in a controlled setting. Compared to Studies I and II, this adds the necessary control regarding available information, distractions, focus on the task and all other benefits that a lab environment provides. The experiment for Study III is fully computer-based and consists of three parts: comprehensive instructions, the main experimental task, and a concluding questionnaire. The course of the experiment is visualized in Section 6 of the Online Appendix.

In the main part of the experiment, participants have to repeatedly decide how much of their endowment to invest in a risky asset. The part of the endowment that is not invested earns no return. As described above, the risky asset is resembled by a binary lottery that has one

positive and one negative outcome.[11] Expected return, standard deviation, and skewness of outcomes is constant for a given asset, but varies between assets to reduce repetitiveness. Each participant sees a total number of 18 different assets in randomized order. For 15 of the 18 assets the participants first observe 15 periods in which a price path develops based on the outcomes of the lottery.[12] In these first 15 periods, they cannot make any investment decisions. Over the subsequent 15 periods (periods 16-30), they make their investment decisions. Immediately after they confirmed their decision, the outcome of this period is realized and the price path continues to develop. All return realizations in the individual periods are independent from one another, and all return properties are clearly communicated to participants.

The price charts for the first 15 periods are preconstructed for 2/3 of the 15 rounds. In these preconstructed scenarios, distributions of period returns are identical but only differ with respect to the order in which returns occur. In preconstructed convex (concave) price charts, negative (positive) returns occur earlier than positive (negative) returns. We included these preconstructed charts to analyze how the price path shape affects the viability of our visual salience measure.[13] Random (non-preconstructed) charts are random realizations of the binary lottery for the first 15 periods, which can result in equal, higher, or lower overall return than in the convex and concave scenarios. For periods 16-30, period returns for all scenarios are truly random outcomes determined by the given parameters of the asset.

Participants use a slider to communicate their investment decisions. To avoid anchoring effects, there is no pre-selected slider position at the beginning of each period. Instead, participants need to click at any position of the slider to activate it; afterwards, they can move

[11] For a detailed specification of the set of binary lotteries, we refer to the Instructions of the experiment in Section 7 of the Online Appendix.

[12] For the three remaining assets, no price path was shown. The analyses for this situation can be obtained from the authors upon request, but does not impact our results but serves as a benchmark w.r.t. the replication of previous results of Nolte and Schneider (2018).

[13] There were no qualitative differences in performance between scenarios, so we do not pursue this analysis further in this paper. Additional data is available from the authors upon request.

the slider position freely to decide on their investment. The current endowment and possible changes in wealth resulting from the chosen investment are depicted on screen to facilitate accessibility of the decision. After the final decision is made and confirmed, a pop-up window informs participants about the realized return and the wealth consequences.

Similar to the design in Study I and II, participants are paid according to one of the 18 decision rounds. This round is determined randomly after the experiment. For this, each participant draws a numbered ping-pong ball out of a concealed urn with 18 balls. This aims at ensuring trust among participants about how their remuneration is determined. Participants then receive a variable payment of 0.001€ for each monetary unit of their final wealth in this round, plus a fixed payment of 8.00€. The experiment concludes with a questionnaire on demographics and self-perceived behavior, the Cognitive Reflection Test (Frederick, 2005), a financial literacy test (FL, Lusardi and Mitchell, 2011), and a risk aversion elicitation task (Holt and Laury, 2002).

We recruited 275 undergraduate and graduate business students at an university in Germany. The experiment was conducted under controlled conditions in the computer labs of the university in 18 separate sessions. On average, it took participants 61 minutes to complete the entire experiment (including instructions and questionnaire), and average remuneration was 16.60€ (about 18.68 USD-$ at the time of the experiment). For comparison, a student research assistant earned 10.00€ per hour at that time.

4.2.2 Results: Study III

To analyze our data, we use all 15 investment decisions of each decision round as our dependent variable. This allows us to incorporate the effect of an unfolding price path in our analyses. Since our asset only has binary outcomes, analyzing the correlation between weights and returns as for the previous studies is not a viable option. Instead, we focus on the CPT value function for visual salience, salience, and CPT weights.

We include expected return, standard deviation, skewness, and actual realizations as control variables. The first three are defined by the characteristics of the risky asset and determine the (normative) attractiveness of the asset. Since we only vary these parameters marginally between decision rounds to prevent repetitiveness, we do not expect to see large effects on any of these measures. Actual realizations are the result of a random process and determine the shape of the respective price path in each round and period. They are operationalized as the average realized returns per period.[14] While expected return, standard deviation, and skewness exclusively account for participants' beliefs, realized returns might influence beliefs and preferences, or both. Participants might believe in autocorrelation (positive realized returns indicate higher future returns) or mean-reversion (positive realized returns indicate lower future returns), ignoring the actual probabilities and outcomes provided. Participants might also act more risk averse (seeking) after positive (negative) realized returns if that pushes the asset in the gain (loss) domain. So while there is no straight-forward expectation we have for the coefficient of this variable, it is important to control for these effects as they would otherwise be captured in our visual salience, salience, or CPT score.

Table 6 reports regression results for each of our three measures individually (Model (1)-(3)) and jointly (Model (4)). We see the same pattern we have observed in Studies I and II with respect to our visual salience measure. Its coefficient is highly significant and has positive explanatory power for investments. In contrast to the first two studies, we see a similar effect both in strength and significance for CPT, while salience remains insignificant.

One drawback of the simplified setup we test in Study III is that we cannot analyze correlations, and therefore our results may be dependent on the exact specification of the value function. We account for this unavoidable deficit with two additional analyses: first, we include

[14]Note that this is synonymous with including the total realized return, realized standard deviation, or total number of periods with a price in-/decrease, since probabilities and outcomes of the binary lottery are constant within each round.

Table 6: Regressions for IA, Study III: CPT Value Function

	(1) IA [%]	(2) IA [%]	(3) IA [%]	(4) IA [%]	(5) IA [%]	(6) IA [%]
Visual Salience	0.0358***			0.0257***	0.0287***	0.0469***
	(0.00814)			(0.00867)	(0.00838)	(0.0119)
CPT		0.0386***		0.0245*	0.0183	0.0163
		(0.0128)		(0.0140)	(0.0153)	(0.0140)
Salience			0.00804	0.00737	0.00524	0.00719
			(0.00766)	(0.00763)	(0.00801)	(0.00794)
Expected return	0.0834	0.0685	0.0778	0.0730	0.0775	0.113*
	(0.0794)	(0.0790)	(0.0796)	(0.0793)	(0.0798)	(0.0601)
Expected Stdv.	-0.00913	0.000798	-0.0242*	0.00271	-0.0261	-0.0409
	(0.0133)	(0.0154)	(0.0129)	(0.0153)	(0.0394)	(0.0270)
Expected skewness	-0.251***	-0.287***	-0.256***	-0.287***	-0.275***	-0.438***
	(0.0410)	(0.0428)	(0.0443)	(0.0457)	(0.0495)	(0.0338)
Average realized return	-0.339***	-0.346***	-0.234***	-0.363***	-0.224	-0.355***
	(0.0427)	(0.0520)	(0.0389)	(0.0547)	(0.167)	(0.120)
CPT equal weights					-0.0488	-0.0665*
					(0.0589)	(0.0403)
Constant	0.115**	0.0940*	0.160***	0.0871	0.172	0.469***
	(0.0478)	(0.0532)	(0.0468)	(0.0530)	(0.119)	(0.114)
Observations	61875	61875	61875	61875	61875	61875
N	275	275	275	275	275	275
R^2	0.024	0.024	0.024	0.024	0.024	

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

The table reports regressions of invested amounts on the Visual Salience model, the CPT model (Barberis et al., 2016) and the Salience model (Bordalo et al., 2012) respectively. Control variables are expected return of the binary lottery, its standard deviation, its skewness, and the average realized return. In Model (5) we also include a CPT value function with equal (1/n) decision weights to control for the effect of value function specification. In Model (6) we use a Tobit regression to account for accumulation of participants who did not invest in that stock. We use fixed effects to control for individual characteristics of participants. All variables are standardized for easy comparison of coefficient sizes and to enhance readability of the table. Robust standard errors clustered on participant level are reported in parentheses.

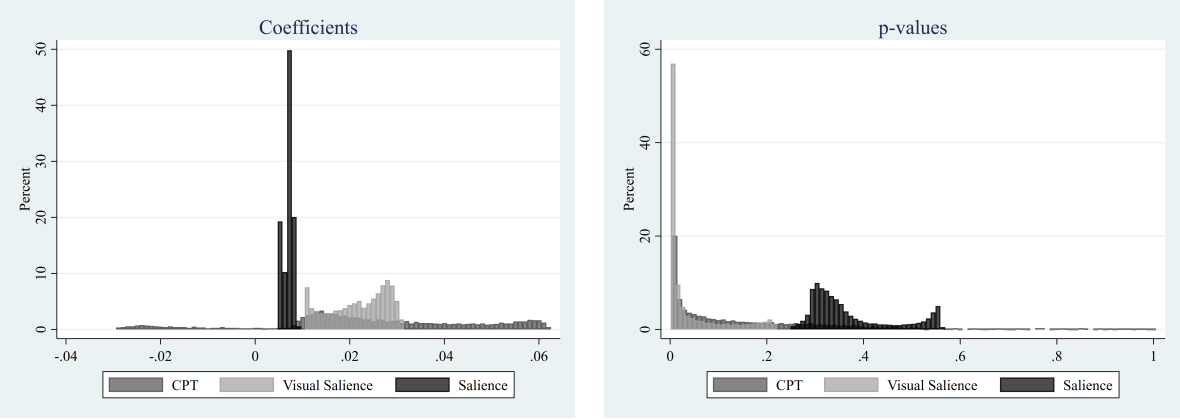
the value of a CPT value function with equal decision weights for each realized return as a control variable. That way, we create a baseline effect that can be attributed to the particular specification of the value function itself. Any explanatory power that the three models retain in this setup has to stem from the decision weights. The effect of including this control variable is depicted in Model (5) of Table 6. The visual salience score retains its significance, and the coefficient remains almost unchanged. CPT on the other hand loses significance as well as strength of relation, indicating that this measure is much more dependent on the exact specification of the value function than visual salience. In Model (6) we also run a Tobit regression to accommodate for accumulation of participants who did not invest in that stock. The results remain qualitatively the same.

For our second analysis to investigate the impact of value function specification, we tested a large number of α-λ combinations. Namely, we varied α between 0.7 and 1 in increments of 0.01, and λ between 0.8 and 3 in increments of 0.05. We then ran Model (4) of Table 6 for each possible alpha-lambda combination within this range. In total, this procedure resulted in 1,350 regression analyses. Figure 7 reports histograms of resulting coefficients for each of our measures on the left hand side, and histograms of p-values on the right hand side. Results confirm that the effect of visual salience is relatively independent of the chosen value function, while the effect of CPT varies strongly with the parameters of the value function. For salience, the choice of value function also doesn't affect results, although coefficients remain small and insignificant.[15] Although there are some parameter constellations where CPT outperforms visual salience in terms of effect strength, we can conclude from Study III that attention driven by visual salience has high explanatory power for investor decision making, even in the simplest of choices and in a highly controlled environment.[16]

[15]Interestingly, we find strongest results for CPT when λ is between 1.2 and 2. This is consistent with results from recent meta studies on loss aversion (Walasek et al., 2018).

[16]In Section 9 of the Online Appendix we also show the parameter sensitivity analysis for the cumulative distribution function of coefficients and p-values.

Figure 7: Parameter sensitivity analysis



The histogram on the left (right) shows the three models' distribution of coefficients (p-values) for all analyzed curvature (α) and loss aversion (λ) combinations. α is varied in increments of 0.01 between 0.7 and 1, λ in increments of 0.05 between 0.8 and 3. In total, the graphs include coefficients and p-values from 1,395 regressions.

5 Conclusion

We explore visual salience as a key determinant of attention in financial decisions. While there are many aspects of a decision making problem that potentially steer attention, human perception is particularly susceptible to visual features and stimuli. In our visual salience model, investors assign decision weights to past returns based on how visually salient the respective area of the corresponding price path is. Put differently, investors tend to pay more attention to and consequently overweigh returns that 'stick out' in some way. Li and Yu (2012) show that investors indeed pay more attention to returns that are near natural anchors, for example the height of a peak (George and Hwang, 2004), the bottom of a trough (Huddart et al., 2009), or during a particularly long streak (Raghubir and Das, 2009).

Although some of these attention-generating returns may be predictable based on statistical moments alone without utilizing a visual attention measure, there is an indefinite number of potential return-price-combinations that could generate more or less visually salient features. Therefore, we rely on modern machine learning techniques to predict salient points in price

paths. Since human perception of visual stimuli is universal and consistent across different domains, we employ an established, pre-trained algorithm to make these saliency predictions. We can confirm the adequacy of this algorithm for price paths by providing evidence that its predictions match actual eye-tracking data closely.

We show in three carefully designed experimental studies that our visual salience model can forecast investment decisions. This holds for real stock price charts, for artificially constructed charts that only vary in the order of the returns but not the returns themselves, and for the easiest possible depiction of price charts in repeated two-outcome lotteries. Compared to other economic models that focus on different aspects of attention and salience like perception of probabilities and return differences, our visual salience model provides higher explanatory power and more accurate investment predictions.

We are cognizant of the fact that investors likely look at a number of additional sources for information before making choices in the real world. Two of our experiments present a stylized setup where the price path is the *only* source they are exposed to, because we wanted to isolate the effect of visual salience. It is likely that when considering these other sources, probability weighting and return differences like in the models of Barberis et al. (2016) and Bordalo et al. (2012) play a larger role in decision-making. As mentioned earlier, we believe that integrating visual salience and recency effects into an integrated model highlighting probability weighting or return differences could deepen our understanding of the underlying mechanisms at play.

Our data clearly shows that economic models benefit from incorporating visual salience features of decision problems. Since visual salience is grounded in basic human visual perception, it is directly measurable via eye-tracking and predictable via algorithms. This allows analyzing the impact of visual salience in other economic domains where visual features of the provided information and alternatives also may play a role. The attention that aspects of the decision problem receive based on their visual salience potentially affects several layers of the decision

making process, for example reference points, outcomes, or even lead to ignoring or considering entire alternatives. Future research will shed further light on these additional applications of visual salience.

References

Abdellaoui, M. (2000), Parameter-free elicitation of utility and probability weighting functions, *Management Science* **46**(11), 1497–1512.

Bali, T. G., Cakici, N. and Whitelaw, R. F. (2011), Maxing out: Stocks as lotteries and the cross–section of expected returns, *Journal of Financial Economics* **99**(2), 427–446.

Barber, B. M. and Odean, T. (2007), All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *The Review of Financial Studies* **21**(2), 785–818.

Barberis, N., Huang, M. and Santos, T. (2001), Prospect theory and asset prices, *Quarterly Journal of Economics* **116**(1), 1–53.

Barberis, N., Mukherjee, A. and Wang, B. (2016), Prospect theory and stock returns: An empirical test, *The Review of Financial Studies* **29**(11), 3068–3107.

Barberis, N., Shleifer, A. and Vishny, R. (1998), A model of investor sentiment, *Journal of Financial Economics* **49**(3), 307–343.

Baucells, M., Weber, M. and Welfens, F. (2011), Reference-point formation and updating, *Management Science* **57**(3), 506–519.

Benartzi, S. and Thaler, R. H. (1995), Myopic loss aversion and the equity premium puzzle, *The Quarterly Journal of Economics* **110**(1), 73–92.

Bordalo, P., Gennaioli, N. and Shleifer, A. (2012), Salience theory of choice under risk, *Quarterly Journal of Economics* **127**(3), 1243–1285.

Bordalo, P., Gennaioli, N. and Shleifer, A. (2019), Memory, attention, and choice, *Technical report*, National Bureau of Economic Research.

Borji, A. and Itti, L. (2015), Cat2000: A large scale fixation dataset for boosting saliency research, *arXiv preprint arXiv:1505.03581*.

Borji, A., Sihite, D. N. and Itti, L. (2012), Quantitative analysis of human-model agreement in visual saliency modeling: A comparative study, *IEEE Transactions on Image Processing* **22**(1), 55–69.

Borji, A., Tavakoli, H. R., Sihite, D. N. and Itti, L. (2013), Analysis of scores, datasets, and models in visual saliency prediction, *Proceedings of the IEEE International Conference on Computer Vision*, pp. 921–928.

Brainard, D. H. (1997), The Psychophysics Toolbox, *Spatial Vision* **10**, 433–436.

Busse, J. A. and Green, T. C. (2002), Market efficiency in real time, *Journal of Financial Economics* **65**(3), 415–437.

Bylinskii, Z., Judd, T., Oliva, A., Torralba, A. and Durand, F. (2018), What do different evaluation metrics tell us about saliency models?, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **41**(3), 740–757.

Camerer, C. F., Johnson, E., Rymon, T. and Sen, S. (1993), Cognition and framing in sequential bargaining for gains and losses, *Frontiers of Game Theory* **104**, 27–47.

Cassella, S. and Gulen, H. (2018), Extrapolation bias and the predictability of stock returns by price-scaled variables, *The Review of Financial Studies* **31**(11), 4345–4397.

Cornia, M., Baraldi, L., Serra, G. and Cucchiara, R. (2018), Predicting Human Eye Fixations via an LSTM-based Saliency Attentive Model, *IEEE Transactions on Image Processing* **27**(10), 5142–5154.

Cosemans, M. and Frehen, R. (2017), Salience theory and stock prices: Empirical evidence, *Working Paper*.

Czaczkes, B. and Ganzach, Y. (1996), The natural selection of prediction heuristics: Anchoring and adjustment versus representativeness, *Journal of Behavioral Decision Making* **9**(2), 125–139.

Da, Z., Engelberg, J. and Gao, P. (2011), In search of attention, *The Journal of Finance* **66**(5), 1461–1499.

DellaVigna, S. and Pope, D. (2017), What motivates effort? Evidence and expert forecasts, *The Review of Economic Studies* **85**(2), 1029–1069.

Diacon, S. and Hasseldine, J. (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.

Diecidue, E., Schmidt, U. and Zank, H. (2009), Parametric weighting functions, *Journal of Economic Theory* **144**(3), 1102–1118.

Fang, L. and Peress, J. (2009), Media coverage and the cross-section of stock returns, *The Journal of Finance* **64**(5), 2023–2052.

Frederick, S. (2005), Cognitive reflection and decision making, *Journal of Economic Perspectives* **19**(4), 25–42.

George, T. J. and Hwang, C.-Y. (2004), The 52-week high and momentum investing, *The Journal of Finance* **59**(5), 2145–2176.

Glaser, M., Iliewa, Z. and Weber, M. (2019), Thinking about prices versus thinking about returns in financial markets, *The Journal of Finance* **74**(6), 2997–3039.

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

Gödker, K. and Lukas, M. (2017), All that glitters is not gold: Asymmetric investor attention in the stock market, *Available at SSRN 3080332*.

Goldstein, W. M. and Einhorn, H. J. (1987), Expression theory and the preference reversal phenomena., *Psychological review* **94**(2), 236.

Grosshans, D. and Zeisberger, S. (2018), All's well that ends well? On the importance of how returns are achieved, *Journal of Banking and Finance* **87**, 397–410.

Hillert, A., Jacobs, H. and Müller, S. (2014), Media makes momentum, *The Review of Financial Studies* **27**(12), 3467–3501.

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

Huber, C. and Huber, J. (2018), Scale matters: Risk perception, return expectations, and investment propensity under different scalings, *Experimental Economics* pp. 1–25.

Huddart, S., Lang, M. and Yetman, M. H. (2009), Volume and price patterns around a stock's 52-week highs and lows: Theory and evidence, *Management Science* **55**(1), 16–31.

Jiang, M., Huang, S., Duan, J. and Zhao, Q. (2015), SALICON: Saliency in Context, *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Judd, T., Durand, F. and Torralba, A. (2012), A benchmark of computational models of saliency to predict human fixations, *Technical Report, MIT*.

Judd, T., Ehinger, K., Durand, F. and Torralba, A. (2009), Learning to predict where humans look, *2009 IEEE 12th International Conference on Computer Vision*, IEEE, pp. 2106–2113.

Kahneman, D. (1979), Prospect theory: An analysis of decisions under risk, *Econometrica* **47**, 278.

Kahneman, D. (2003), Maps of bounded rationality: Psychology for behavioral economics, *American Economic Review* **93**(5), 1449–1475.

Kahneman, D. and Tversky, A. (1972), Subjective probability: A judgment of representativeness, *Cognitive Psychology* **3**(3), 430–454.

Kahneman, D. and Tversky, A. (1973), On the psychology of prediction., *Psychological Review* **80**(4), 237.

Kennedy, R., Clifford, S., Burleigh, T., Waggoner, P. and Jewell, R. (2018), The shape of and solutions to the MTurk quality crisis, *Available at SSRN*.

Koester, A., Lundholm, R. and Soliman, M. (2016), Attracting attention in a limited attention world: exploring the causes and consequences of extreme positive earnings surprises, *Management Science* **62**(10), 2871–2896.

Krajbich, I. and Rangel, A. (2011), Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions, *Proceedings of the National Academy of Sciences* **108**(33), 13852–13857.

Krajbich, I., Armel, C. and Rangel, A. (2010), Visual fixations and the computation and comparison of value in simple choice, *Nature Neuroscience* **13**(10), 1292.

Kumar, A. (2009), Who gambles in the stock market?, *The Journal of Finance* **64**(4), 1889–1933.

Kumar, A., Ruenzi, S. and Ungeheuer, M. (2019), Daily winners and losers, *Available at SSRN 2931545*.

Le Meur, O. and Baccino, T. (2013), Methods for comparing scanpaths and saliency maps: strengths and weaknesses, *Behavior Research Methods* **45**(1), 251–266.

Li, J. and Yu, J. (2012), Investor attention, psychological anchors, and stock return predictability, *Journal of Financial Economics* **104**(2), 401–419.

Li, X. and Camerer, C. (2019), Using Visual Salience in Empirical Game Theory, *Available at SSRN 3308886*.

Lusardi, A. and Mitchell, O. S. (2011), Financial literacy around the world: An overview, *Journal of Pension Economics and Finance* **10**(4), 497–508.

Mohrschladt, H. and Schneider, J. C. (2018), The Idiosyncratic Volatility Puzzle and its Interplay with Sophisticated and Private Investors, *Available at SSRN 3131703*.

Mormann, M. M. and Frydman, C. (2017), The role of salience and attention in choice under risk: An experimental investigation, *Working Paper*.

Mullett, T. L. and Stewart, N. (2016), Implications of visual attention phenomena for models of preferential choice., *Decision* **3**(4), 231.

Nolte, S. and Schneider, J. C. (2018), How price path characteristics shape investment behavior, *Journal of Economic Behavior and Organization* **154**, 33–59.

Odean, T. (1999), Do investors trade too much?, *American Economic Review* **89**(5), 1279–1298.

Pelli, D. G. (1997), The VideoToolbox software for visual psychophysics: transforming numbers into movies, *Spatial Vision* **10**, 437–442.

Peters, R. J., Iyer, A., Itti, L. and Koch, C. (2005), Components of bottom-up gaze allocation in natural images, *Vision Research* **45**(18), 2397–2416.

Prelec, D. (1998), The probability weighting function, *Econometrica* **66**(3), 497–527.

Raghubir, P. and Das, S. R. (2009), The long and short of it: why are stocks with shorter runs preferred?, *Journal of Consumer Research* **36**(6), 964–982.

Ross, J., Irani, L., Silberman, M., Zaldivar, A. and Tomlinson, B. (2010), Who are the crowdworkers?: shifting demographics in mechanical turk, *CHI'10 Extended Abstracts on Human factors in Computing Systems*, ACM, pp. 2863–2872.

Towal, R. B., Mormann, M. and Koch, C. (2013), Simultaneous modeling of visual saliency and value computation improves predictions of economic choice, *Proceedings of the National Academy of Sciences* **110**(40), E3858–E3867.

Tversky, A. and Kahneman, D. (1992), Cumulative prospect theory: An analysis of decision under uncertainty, *Journal of Risk and Uncertainty* **5**(4), 297–323.

Ungeheuer, M. (2017), Stock returns and the cross-section of investor attention, *Available at SSRN 2931547*.

Wakker, P. P. (2010), *Prospect theory: For risk and ambiguity*, Cambridge university press.

Walasek, L., Mullett, T. L. and Stewart, N. (2018), A meta-analysis of loss aversion in risky contexts, *Working Paper*.

Weber, E. U., Siebenmorgen, N. and Weber, M. (2005), Communicating asset risk: How name recognition and the format of historic volatility information affect risk perception and investment decisions, *Risk Analysis: An International Journal* **25**(3), 597–609.

Wu, G. and Gonzalez, R. (1996), Curvature of the probability weighting function, *Management Science* **42**(12), 1676–1690.