Decision Weights for Experimental Asset Prices Based on Visual Salience

Devdeepa Bose  
California Institute of Technology, USA

Henning Cordes  
University of Münster, Germany

Sven Nolte  
Institute for Management Research, Radboud University Nijmegen, The Netherlands

Judith Christiane Schneider  
Institute for Risk and Insurance, Leibniz University Hannover, Germany

Colin Farrell Camerer  
California Institute of Technology, USA

We apply a machine-learning algorithm, calibrated using general human vision, to predict the visual salience of prices of stock price charts. We hypothesize that the visual salience of adjacent prices increases the decision weights on returns computed from those prices. We analyze the inferred impact of these weights in two experimental studies that use either historical price charts or simpler artificial sequences. We find that decision weights derived from visual salience are associated with experimental investments. The predictability is not subsumed by statistical features and goes beyond established models. (JEL G12, G40)

Received July 8, 2020; editorial decision October 15, 2021 by Editor Tarun Ramadorai. Authors have furnished an Internet Appendix, which is available on the Oxford University Press Web site next to the link to the final published paper online.

We thank Nicholas Barberis, Lawrence Jin, Alex Imas, Elise Payzan-LeNestour, Cindy Xiong, and Thomas Langer and seminar and conference participants at Warwick Business School, Radboud University, University of Münster, Caltech, the 2019 Southwestern Economic and Behavioral Economics Conference, the 2019 Experimental Science Association, the 2019 Mid-Atlantic Meeting on Behavioral and Experimental Economics, the 2019 Boulder Summer Conference on Consumer Financial Decision Making, and the 2020 NBER Behavioral Finance Working Group Summer Meeting for helpful comments and suggestions. Financial support was provided by the Tianqiao and Chrissy Chen Center for Social and Decision Neuroscience Leadership Chair, Behavioral and Neuroeconomics Discovery Fund, the Alfred P. Sloan Foundation, and the Ronald and Maxine Linde Institute for Economics and Management Sciences. Supplementary data can be found on The Review of Financial Studies web site. Send correspondence to Sven Nolte, sven.nolte@ru.nl.
Attention is a scarce resource. High-attention events are therefore thought to affect investment decisions, and evidence for this attention hypothesis can be found in finance (Barber and Odean 2007; Hillert, Jacobs, and Müller 2014; Mohrschlatt and Schneider 2021). We take this hypothesis in a new direction.

An immense amount of information on a particular stock confronts the investor who visits the website of an online broker, a newspaper, or other financial magazine. One ubiquitous source of financial information is the historical price chart, which, according to the field data in Glaser, Iliewa, and Weber (2019), is presented by the vast majority of financial information providers (above 90%). That the visual properties of these charts steer investor attention is reflected in investment decisions (Grosshans and Zeisberger 2018; Nolte and Schneider 2018; Glaser, Iliewa, and Weber 2019). More generally, humans are naturally drawn to images, and trying to shift focus away from such visualizations incurs significant mental costs (Sanchez and Wiley 2006).

We add a novel concept to understand the impact of this information: bottom-up visual attention. We propose that initial attention toward 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). Many subsequent studies have shown that visual fixations can predict economic choices in different domains, such as food choice (Grebitus, Roosen, and Seitz 2015), and that manipulating fixation lengths can bias decisions Armel, Beaumel, and Rangel (2008).

Our analysis employs one of the latest in a steadily improving sequence of algorithms incorporating neuroscientific properties to predict what people look at in visual images. This algorithm (Saliency Attentive Model [SAM] from Cornia et al. 2018) is “trained,” in machine learning terms, on actual eye-tracking data from several thousand participants freely gazing at many different images.

We incorporate visual salience into a preference model to predict investment choices. Following Barberis, Mukherjee, and Wang (2016) and Cosemans and Frehen (2021), we assume that investors mentally represent the distribution of a stock’s past returns as a proxy for future returns. While the objective probabilities of past returns are known (as they have been realized), they will be distorted by the visual attention paid to specific stock returns. The central assumption is that outcomes in past dates serve as a proxy for outcomes in future states. The distorted probabilities (decision weights) and the outcomes are then evaluated by a reasonable preference functional, for example, gain-loss utility.

To gain evidence for how useful SAM is in predicting visual attention in price charts (independent of top-down investment goals), we conducted an experiment with eye-tracking where participants are exposed to a sample of
Decision weights for experimental asset prices

price charts, and are motivated to pay attention by a later memory test. We compare participants’ fixations with SAM predictions, calculating the fit with four metrics established to evaluate saliency models (Le Meur and Baccino 2013; Bylinskii et al. 2018). We conclude that SAM, although not trained on stock price path images, is associated with what people perceive as salient in a price path.

It is also important to understand whether visual attention weights generated by SAM are associated with statistical features of the historical return distribution. This ties into the more general discussion of “explainable AI” to create more transparent foundations for algorithms predicting human perception or choice (Hinton, Vinyals, and Dean 2015; Lipton 2018; Ras, Gerven, and Haselager 2018; Arrieta et al. 2020; Belle and Papantonis 2021; Fan, Xiong, and Wang 2021). More specifically, we utilize a popular method for creating explainability called “feature relevance.” We regress SAM weights on 29 statistical measures (features) from the asset pricing literature, reduced into fewer components by principal component analysis (PCA). This analysis suggests that a maximum of 20% of the variance in SAM decision weights can be explained by coded statistical features. Two groups of features (i.e., PCA features) that stand out are linked to price levels and return variability.

Finally, we ask whether the visual salience model can add explanatory power when predicting investments beyond these statistical features, and beyond existing preference models like cumulative prospect theory (CPT) (Tversky and Kahneman 1992) as applied by Barberis, Mukherjee, and Wang (2016), or salience theory (Bordalo, Gennaioli, and Shleifer 2012; 2013) as applied by Cosemans and Frehen (2021). We conduct two experimental studies to evaluate the performance of the visual salience model.

Study 1 uses empirical price charts drawn from actual stock returns. Experimental investors see these charts, decide how much to invest in the respective asset, and earn money based on their investment allocations, and actual future returns of the asset. A model where decision weights are based on visual salience of the historical prices can partially explain their investments.

Study 2 steps back from historical prices by creating the simplest possible depiction of price charts: artificial, repeated binary lotteries. This enables us to control for many factors that could influence investment decisions like image or task complexity, as well as the environment, attention, and knowledge of participants. We show that the visual salience model is also able to predict future investments in this design.

We conclude our paper by presenting two further implications of visual salience for the investment decision process. First, we discuss how bottom-up visual salience could be unified with theories describing the top-down processing part of an investment decision. While bottom-up processing assembles and integrates the sensory information presented in a price path, top-down processing interprets this information based on goal-based models, ideas, and previous experience. Most (normative or descriptive) models of
financial decision-making have an implicit assumption of deliberate, top-down guided agents. We postulate that decision weights as in Barberis, Mukherjee, and Wang (2016) and salience theory (Bordalo, Gennaioli, and Shleifer 2012; 2013) as applied by Cosemans and Frehen (2021) capture top-down information processing. Our results indicate that a model combining these top-down and our bottom-up decision weights improves the predictability of investment choices.

Second, we analyze how visual salience relates to the formation of expectations about future stock returns. In a first step, we formulate a model of visual salience-based belief formation and test its predictive power for actual beliefs about future risk and return. The outcome of the model shows that the influence is low for predicting return expectations, and moderate for risk expectations. In a second step, we confirm that our visual salience model retains its explanatory power even when controlling for participants’ self-stated risk and return expectations. This implies that visual salience captures an aspect of the decision problem affecting investment choices that goes beyond beliefs about the first two moments of the stock return distributions.

Our research connects to four strands of literature in finance and economics. We briefly summarize these here and in more detail in Section 1 of the Internet Appendix.

First, several studies explore how the framing or presentation of historical returns activates different biases and heuristics investors employ, and influences their decision-making (Weber, Siebenmorgen, and Weber 2005; Diacon and Hasseldine 2007; Glaser et al. 2007; Grosshans and Zeisberger 2018; Nolte and Schneider 2018; Glaser, Iliewa, and Weber 2019). We show that presentation formats influence investments through their visual salience. Second, the large amount of information being effortlessly available implies that investors (both retail and institutional), cannot pay attention to everything. Many studies show that heightened attention influences investment choices (Odean 1999; Busse and Green 2002; Barber and Odean 2007; Kumar 2009; Da, Engelberg, and Gao 2011; Ungeheuer 2017). We relate to this by predicting (and also measuring) visual attention in investment experiments. Third, behavioral models propose different ways in which past returns are weighted and integrated to form expectations about future returns and drive investment (Barberis, Mukherjee, and Wang 2016; Cosemans and Frehen 2021). We suggest a new weighting scheme for these models, using visual salience to weight past returns. Fourth, visual salience also has been used to explain choices in a couple of studies about consumer and strategic game choices (Towal and Mommam 2013; Li and Camerer 2021). We extend the range of applications to financial investment decisions.

1. Theoretical Framework

Our key assumption, following Barberis, Mukherjee, and Wang (2016) and Cosemans and Frehen (2021), is that investors mentally represent a stock based on its past returns, as they are an omnipresent and easily accessible
Decision weights for experimental asset prices

proxy for a stock’s future return distribution.¹ As pointed out earlier, most financial information providers depict a historical price chart. We therefore conjecture that such a representation is likely to come easily to an investor’s mind. Let $X$ be a random variable representing stock returns, with past realizations $x_1, \ldots, x_m$ over a specific horizon. The dates of the past realizations serve as a proxy for the state space, which is unknown to the investor, and describes the choice problem under risk. Barberis, Shleifer, and Vishny (1998) argue that such an approximation is psychologically rooted in the “representativeness heuristic,” a concept introduced by Kahneman and Tversky (1972, 1973). This heuristic leads the investor to think that the limited sample of past realizations is representative of the full distribution of possible (future) outcomes.

The historical return distribution assigns an equal probability to each realization. The investor, however, does not use these objective probabilities to evaluate the stock but is hypothesized to transform them into decision weights during the evaluation process. We use the visual salience/attention paid to specific stock returns to transform the objective probabilities into decision weights. The idea to weight specific components of the decision calculus with the level of attention they receive can be generalized to any finance model, and is not restricted to returns.² We assume the most common representation—returns—in our framework to be comparable to the vast majority of finance research.

To obtain decision weights based on visual salience, price charts are inputs into SAM, the visual salience algorithm. SAM then outputs a greyscale image assigning higher values (close to one) to pixels that contain features that are more visually salient and lower values (close to zero) to pixels that contain features that are less visually salient. In the end, every pixel of a price chart has a salience value from 0 to 1, which comes from the SAM algorithm. We aggregate the visual salience values for all the pixels associated with each price, and average the values across consecutive prices to arrive at a salience value for each return.³ Objective probabilities are weighted by those predicted salience

¹ Returns are the typical means of analysis in finance. However, from a visual perspective, price differences are easier to detect and might enter the decision calculus instead. In Internet Appendix Section 8, we apply this reasoning to the data and run all our analyses with price differences instead of returns. We show that our results are robust with respect to this assumption.

² One could, for example, also think about a model in which investors mentally represent a stock based on its past prices instead of returns or price differences. In the spirit of our framework, these prices would be weighted by the attention they receive, which would be captured by visual salience of the respective prices in the price chart. When testing this simple attention-weighted prices model in Internet Appendix Section 8, we find that it retains significant predictive power for investments, even after controlling for the benchmark of equally weighted prices.

³ We identify the pixel where each price appears on the chart and then draw a circle with a one pixel radius around it. We then aggregate the values of all pixels in this circle to obtain the visual salience values associated with each price. In Internet Appendix Section 5, we test our results with respect to these assumptions. We run our analyses for different radii (0.5 pixels, 1.5 pixels) around each price, and for any combination between previous and subsequent price for return salience. Our results remain qualitatively unchanged.
values, and normalized to sum to one. In a sense, a hidden “parameterization” generates the salience value, coming from the complex convolutional “deep learning” SAM neural network that has been “trained” (in the machine learning sense) on an entirely different set of data from people looking at images (which are not the images of price charts we use). The accuracy of this method is therefore likely to be a lower bound on how well more specialized neural networks would predict if they were trained based on what people actually look at in price charts.

Using these decision weights, we now propose a formal model. We assume that an investor values a stock by using a combination of past returns and decision weights. Thus, she evaluates $X$ according to:

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

(1)

where $v(\cdot)$ is a continuous, strictly increasing value function, and $\pi_k$ represents the visual salience decision weight associated with the historical realization mapped into a future state with associated probability $p_k$. Then, $\pi_k$ is defined as

$$\pi_{VS} = p_k \times l_k, \sum_{k} \pi_{VS} = 1, \text{ with the visual salience weight } l_k = \frac{\phi_k}{\sum_i \phi_i p_i}.$$  

(2)

and $\phi_k = \frac{SAM(P_{k-1}) + SAM(P_k)}{2}$, where $SAM(P_k)$ denotes the visual salience of price $P_k$ as predicted by SAM.

It is crucial to note that this weighting function does not model beliefs or expectations, but only how beliefs (in this case, probabilities) enter a decision calculus. For example, in the CPT framework, a lottery participant that has a 1% chance of winning, could say that she believes the probability of winning to be 1%, but then treat that chance as more similar to 5% in her decision to participate. Similarly, for (visual) salience theory in our application, returns that are more (visually) salient are not necessarily actually considered to be more likely to occur, but we assume that they are treated in the investment decision as if they were. So our decision weights do not try to simulate what investors actually believe, but rather model how beliefs (in our case objective probabilities or, rather, empirical frequencies) appear to be distorted in order to closely match actual investment choices. Whether visual salience (also or exclusively) captures beliefs is an interesting follow-up question, but goes beyond our simpler, first-step question of how the visualization of price charts affects investment choices. We further investigate and discuss the connection between visual salience and belief formation in Section 5.2.

---

4 Recent research in neuroeconomics motivates the use of attention weights by showing that subjects in the laboratory attach more weight to attributes that they attend to more when making binary or multiattribute choices, as well as choices among risky lotteries (Krajbich, Armel, and Rangel 2010; Krajbich and Rangel 2011; Frydman and Mormann 2017; Mullet and Stewart 2016).

5 We assume $p_k$ to be equal to the empirical frequency of the respective return.
Decision weights for experimental asset prices

Figure 1
Visual salience
This figure shows three price paths 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 resultant visual salience weights are depicted below each graph.

An example will illustrate how visual salience determines decision weights. Figure 1 depicts three 15-period simulated price paths. The three paths have the same 15 returns that are shuffled into three different temporal orders. The empirical frequencies for each of the 15 returns are equal, \( p_k = 1/15, k \in \{1, \ldots, 15\} \), and these \( p_k \) will be treated as probabilities of the respective returns that are weighted by visual salience decision weights in the decision process.

To determine visual salience decision weights, we input price charts into the SAM algorithm. We aggregate the visual salience values as described above to obtain a decision weight for each return. This is illustrated by the overlay of the heatmap on the different price paths, with richer (darker) colors corresponding with higher salience.

To see how visual salience changes, look first at the path on the left. The highest return is the first one, but it receives the lowest visual salience weight. Instead, from previous samplings, SAM predicts that the eye will be drawn to the second and third returns, both of which have high salience weights (we will

---

6. The paths were generated by an AR(1) process creating prices \( x_t \), with \( x_{t+1} = \rho x_t + \epsilon_t \), and \( \rho = 0.6 \).
return to the fascinating and difficult question of what statistical features are predicted to be visually salient later). The bottom panel shows that the visual salience weights for the second and third returns, as well as the last two returns, get higher weights.

Things change for the price path in the middle. Now the highest return is the second-to-last return in the time series; and as the heatmap and weighting charts show, it receives the highest of all the salience weights. The identical highest return received a low visual salience weight in the first price chart.

The two examples show that the very same absolute return can shift from being least salient to most salient, depending on its location in the price series. Finally, in the chart on the right, the highest return is the seventh return, and it receives a salience weight that is just below average.

2. Saliency Attentive Model

When investors look at a price chart, their gaze is drawn toward visually salient regions. But what exactly is visual salience? A long, impressively accumulating tradition of computational neuroscientists use facts about actual brain processes to emulate human salience in the form of models that predict likely human gaze locations for a wide range of natural images. One of the earliest models is Itti and Koch (2000). They used evidence about coarse to fine encoding in layers of visual cortex and salience of color, contrast, and orientation, to construct a computational process analogous to how the brain works. The algorithms take natural images as inputs and outputs of predicted salience. Their pioneering model, and the many upgrades that followed in rapid order in computer vision, are one of the most impressive examples of artificially constructing machine systems to reproduce human performance.

We use one such model created by the Aimage Lab at the University of Modena and Reggio Emilia called the Saliency Attentive Model (SAM) (Cornia et al. 2018). SAM is a neural network with three main blocks: a feature extraction dilated 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 block, which calibrates how focused human gazes are to the center of images.

2.1 Off-the-shelf SAM

SAM has been trained and validated on a combination of four image data sets: SALICON (Jiang et al. 2015), MIT300 (Judd, Durand, and Torralba 2012), MIT1003 (Judd et al. 2009), and CAT2000 (Borji and Itti 2015). Together, these data sets include more than 23,000 images that display complex everyday scenes
containing common objects in their natural context. Sources for these images include Microsoft COCO, Flickr, and LabelMe. These images 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. SAM makes predictions about portions of an image that will be fixated on within the first few seconds of gaze. We hypothesize that these initial seconds are enough for agents to form a heuristic about potential future returns when shown images of historical price charts. This is very similar to the ideas outlined in Barberis, Mukherjee, and Wang (2016).

SAM identifies features of the images that correlate with a higher number of fixations made by human observers. The identified features could be low level, such as horizontal and vertical lines, or brighter colors, or shifts in contrast and shapes, but also could be high level like faces, trees or animals (or statistical features of price paths like in our studies). 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. 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 more likely to be ignored.

It is well-established in visual neuroscience that attention is guided by a combination of universal, rapid, so-called “bottom-up” features, and so-called “top-down” processes. The two processes are also called “stimulus driven” and “goal directed.” Top-down processes are derived from special experiences and goals that will guide attention. For example, both familiar and novel features will grab attention. Goals will influence attention too. A person first entering a party may be looking for a friend, for an enemy to avoid, where the bar is, where the bathroom is, or whether a man is wearing a wedding band. These scenarios can be described as top-down influences on attention caused by different goals.

Similarly, an investor who is deciding whether to sell may look at different features of a price path than an investor deciding whether to buy (Nolte and Schneider 2018). Because SAM is trained on a large set of natural images,
where people freely gaze, it predicts visual search that is most universal and independent of specific top-down goals, that is, bottom-up salience. Since investors, including participants in our experiments, do have goals, visual salience predicted by SAM is likely to be either a poor prediction of how they invest or a lower bound on how well algorithms customized to predict investment could perform. Nonetheless, given that bottom-up salience is an automatic, cost-free, and involuntary process, it is highly likely that it will retain at least some effect on the final decision, regardless of the type and quality of associated top-down goals. Note that Li and Camerer (2021) have used SAM to successfully predict choices in experimental games where people select locations in natural images, trying to match or mismatch their partners or opponents. They also showed that SAM salience can explain mistakes in choosing low-value objects because they are salient. Their results show that SAM can be used to predict a behavior of economic interest, which it was not specially designed for.

2.2 SAM and price paths

SAM is trained on images that span a variety of domains, because the algorithm is intended to capture general features of what is visually salient during perception of a huge variety of images people see every day. Our approach tests the rather bold hypothesis that decision weights derived from general visual salience also might be predictive in a specialized domain, such as financial price paths.

To gain evidence for how useful SAM is in predicting visual attention in price charts (independent of top-down investment goals), we conducted an eye-tracking experiment where participants are exposed to a sample of price charts. This is an important step to provide a proof of concept, which is by no means trivial or redundant. SAM was trained on natural images, not on price charts. If the participants’ fixations do not match with SAM predictions, off-the-shelf SAM has little hope of predicting investment choices.

We recorded $N=57$ Caltech student participants’ eye-movements and fixations, while they looked at price paths on a computer monitor. To parallel the procedure originally used with MIT300 images (Judd, Durand, and Torralba 2012), we provided participants 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.” This task is called “free gaze” because the subjects are not asked to make any choices based on what they see. The memory test afterward is simply a device to guarantee that they are paying attention.

Images were only shown for three seconds each so that eye fixations would reflect the parts to which initial attention is allocated. Participants were
Decision weights for experimental asset prices

presented with a total of 60 price path images. Thirty of these paths were selected to match those in study 2 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 1.

The participants were tested individually using an EyeLink 1000 Plus (SR Research, Osgoode, Ontario, Canada) eye-tracker. Monocular eye movements were recorded at 500 Hz, 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 proportionally 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. One degree of visual angle is typically used both (1) as an estimate of the size of the human fovea (e.g., how much of the image a participant has in focus during a fixation) and (2) to account for measurement error in the eye tracking setup. This measure is identical to what is used to analyze images in the MIT300 data set (Bylinskii et al. 2018), one of the data sets on which the SAM algorithm was trained.

Figure 2 gives two examples of the eye-tracking data (called “groundtruth density maps”) from the experiment on the left, and the SAM-predicted heatmaps overlaid with the price path participants looked at on the right. Simple visual inspection tells us that fixations and predictions overlap substantially. We measure the similarity between the eye-tracking fixation data from the experimental subjects and the SAM predictions by calculating four visual salience metrics that help quantitatively evaluate how well SAM predicts salient points within price paths. For an overview of suitable metrics, we refer to Le Meur and Baccino (2013) and Bylinskii et al. (2018). The first metric belongs to a group of measures 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) provides the degree of similarity of two salience maps. A value of 1.0 indicates perfect classification, while a value of 0.5 indicates random chance classification. We also include the Normalized Scanpath Saliency (NSS), which is a simple correspondence measure between saliency maps and ground truth, computed as the average normalized saliency at fixated locations. Chance is at zero, a positive NSS indicates correspondence between maps above
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.

Chance, and negative NSS indicates anticorrespondence. For instance, a score of one corresponds to fixations falling on portions of the saliency map with a saliency value one standard deviation above average. The simplest measure we include is the Pearson correlation coefficient (CC). Lastly, as a measure of dissimilarity, we take the Kullback-Leibler (KL) divergence, which is a measure from information theory, that captures the overall dissimilarity between two distributions.

Table 1 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’ fixations maps with heatmaps where (for each price path image) fixations were randomly assigned. SAM’s performance on price paths is much closer to its performance on domain-neutral images than it is to randomly assigned fixations, for all our metrics. We conclude that SAM, although not originally trained on human gaze at price path images, captures well what people perceive as salient within a price path, and serves as a good lower bound for predicting gaze in price charts.
Decision weights for experimental asset prices

Table 1
Fixation data and evaluation metrics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>p5</th>
<th>p25</th>
<th>p50</th>
<th>p75</th>
<th>p95</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of fixations</td>
<td>13.04</td>
<td>6.01</td>
<td>7</td>
<td>10</td>
<td>12</td>
<td>14</td>
<td>26</td>
</tr>
<tr>
<td>Duration of fixations (ms)</td>
<td>344.76</td>
<td>441.30</td>
<td>94</td>
<td>172</td>
<td>238</td>
<td>370</td>
<td>840</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Evaluation metric</th>
<th>AUC</th>
<th>NSS</th>
<th>CC</th>
<th>KL</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAM (original domain-neutral)</td>
<td>0.87</td>
<td>2.34</td>
<td>0.78</td>
<td>1.27</td>
</tr>
<tr>
<td>SAM vs. fixations (price paths)</td>
<td>0.81</td>
<td>1.17</td>
<td>0.52</td>
<td>1.58</td>
</tr>
<tr>
<td>Random vs. fixations (price paths)</td>
<td>0.50</td>
<td>0.002</td>
<td>0.07</td>
<td>12.57</td>
</tr>
<tr>
<td>Range (worst-best)</td>
<td>[0,1]</td>
<td>[−∞, 0]</td>
<td>[−1, 1]</td>
<td>[1, ∞]</td>
</tr>
</tbody>
</table>

The top part of the table reports summary statistics including quantiles (e.g. p5 is 5%) regarding the number of fixations for each trial and the fixation duration (in milliseconds). Each trial consists of one of our 57 subjects looking at one of a total of 60 price paths, so there were 571×60=3,420 trials in total. The bottom part of the table reports four evaluation metrics. We compare SAM’s performance on the domain-neutral images it was trained on with SAM’s performance on human eye fixations for price path images. We also compute metrics for randomly assigned fixations to provide an intuitive lower bound for SAM’s performance on price path images. We include one metric using receiver operating characteristics analysis and calculating the area under the curve: AUC (Judd et al. 2009). We also include the Normalized Scanpath Saliency (NSS), Pearson Correlation Coefficient (CC), and Kullback-Leibler divergence (KL).

2.3 SAM and statistical features of price paths

The previous section showed that off-the-shelf SAM has the ability to predict visually salient features of price paths. This justifies our ultimate goal of using attention paid to salient features as predicted by SAM to model meaningful decision weights.

Before we can fully test this conjecture, we have to answer the question whether SAM merely captures similar effects as statistical features of price paths. While we cannot directly identify the “features” in SAM’s neural network, it is possible that some of the features are similar to the statistical features of the return distribution in these price charts. For example, low-level visual attention is known to be captured by differences in orientation, for example, in a picture of a forest, attention is immediately directed to a tree that is bent over. A certain kind of return in a price series might have a similar property of unusual orientation—slope, similar to nonvertical angle—that attracts attention.

Note that SAM is designed to capture rapid, implicit attention to “bottom-up” salient parts of a price path. Many statistical features are instead related to active and higher-level cognitive “top-down” calculations and would seem, a priori, to be poorly noticed by rapid attention. Bottom-up salience of prices or price patterns increases the more visually distinct they are (Wolfe and Horowitz 2017). Statistical features of a price path likely relate to these distinct prices or price patterns. We therefore select common statistical features from the asset pricing literature, designed to capture relevant properties of the return distribution and time series like autocorrelation, dispersion, or jumps (e.g., George and Hwang 2004; Mizraich and Weerts 2009; Bali, Cakici, and Whitelaw 2011; Chen, Hong, and Stein 2001; Chabi, Ruenzi, and Weigert 2018; Raghubir and Das 2010; McLean 2010; Stambaugh, Yu, and Yuan 2015;
Jegadeesh and Titman 1993; Conrad, Kapadia, and Xing 2014). As stated above, these features do not necessarily have a direct link to visual salience but can be related to bottom-up or top-down visual attention. Nevertheless, to provide an intuitive connection between attention and statistical asset pricing features, we categorize the features according to three dimensions known to capture visual attention in general and that are most likely related to price paths (from Wolfe and Horowitz 2017). Extra attention is paid to visual stimuli that have distinct orientation, size, or curvature (elements of shape; Itti and Koch 2000). Additionally, areas in an image that are new, unfamiliar, or break patterns (creating novelty) and those that contrast with a group (contrast) are also known to grab attention (Wolfe and Horowitz 2017). We restrict the categorization of statistical features to one category per feature, even though many of the features likely capture several categories. For example, a negative price drop will be surprising to an investor (so we therefore categorize it as novel). Such a price drop will also affect the visual appearance (shape) of the price path, and will typically stand out (contrasts) compared to all other returns of the path.

Section 3 of our Internet Appendix describes the methodology and the chosen measures in detail.

In total we consider 29 statistical features and run a simple “feature relevance” analysis. We analyze pairwise correlations between a long list of candidate statistical measures and SAM weights in a first step. Then, we regress SAM weights on common features of the statistical measures identified by principal component analysis (PCA).

These analyses suggest that a maximum of 20% of the variance in SAM decision weights can be explained by coded statistical features. One group of features (i.e., one principal component) describes price levels, while another describes variability. Then a long list of components mostly weight a single statistical feature (e.g., high-magnitude jumps).

Obviously, this analysis does not encompass every link between visual salience and statistical features. In particular, there is a link to technical analysis, although that approach generally hypothesizes abstract—and often slippery—definitions of “high-level” features, such as a “resistance level” that may not correspond to anything visually regular. However, further exploration would remove focus from the main goal of our paper, which is to describe how visual salience can be applied to price charts, and show tentative evidence associated with experimental investment decisions.

We conclude that while some features of the return distribution are correlated with visual salience (the 20% of explained variance), the SAM algorithm appears to mostly capture properties of the price path that cannot be replicated by merely combining simple statistics. To fully account for this fact, we control for all statistical features included in this section when trying to explain investment behavior (and find that visual salience has significant explanatory power beyond these statistical features).
Table 2
Experimental studies summary

<table>
<thead>
<tr>
<th></th>
<th>Study 1</th>
<th>Study 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective</td>
<td>Test VS in real price paths</td>
<td>Test VS in simplified setting</td>
</tr>
<tr>
<td>Platform</td>
<td>MTurk</td>
<td>Laboratory</td>
</tr>
<tr>
<td>Price path types</td>
<td>Empirical daily, 1-year period, static (CRSP 2017)</td>
<td>Binary outcome charts, dynamic (15 observation periods 15 investment periods) 2/3 predetermined, 1/3 random</td>
</tr>
<tr>
<td># of participants</td>
<td>500</td>
<td>275</td>
</tr>
<tr>
<td># of price paths</td>
<td>1,000 stocks (each evaluated by four participants)</td>
<td>4,950 unique charts (18 charts evaluated by each participant)</td>
</tr>
</tbody>
</table>

The table summarizes the two experimental studies analyzed to explore the efficacy of visual salience in predicting investment levels.

3. Experiments and Empirical Strategy

Two experiments were conducted to evaluate whether visual salience helps predict investment choices. Table 2 collects the key features of the experiments.

Study 1 is run on Amazon Mechanical Turk (MTurk). Study 2 is a lab experiment at a German university with student participants. 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.

Study 1 used experimental investments based on price charts similar to charts investors often see in the field (constructed from actual daily CRSP returns over a 1-year period). Study 2 considers sequential binary lotteries, the simplest possible depiction of price charts controlling for many factors that could influence investment decisions, like image and task complexity. Subjects are also told the data generating process behind the price paths such that an optimal choice can be computed. Thus, any effect of visual salience in this study highlights the idea that human decision-making cannot easily circumvent the impact of how information is visualized. Moreover, we have perfect control over the environment and attention of the participants.\(^{10}\)

3.1 MTurk experiment

For study 1, we recruited 500 participants from MTurk (Goodman, Cryder, and Cheema 2013). Participants earned a US$2 fixed fee and a variable completion fee, the latter depending on their choices within the experiment. The variable part was, on average, US$0.94. Completion times averaged 12 minutes (with 95% of completion times between 5 and 30 minutes).

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.

\(^{10}\) In both studies, the start price is fixed at 100 monetary units to keep the scaling of the y-axis constant and to limit psychological framing effects (e.g., Glaser et al. 2007; Huber and Huber 2018).
We show each participant eight different, preselected, 1-year price charts. The 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 1. 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 1,000 monetary units into this stock vs. a safe bank account.\textsuperscript{11} We emphasize in the experimental instructions 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 also can be much higher or lower. For each price chart, participants make their investment decisions by using a slider, which allows them to invest every amount between 0\% and 100\% of their initial endowment. To avoid anchoring effects, we do not place a preselected slider position at the beginning of each period. Instead, participants need to click at any position of the slider to activate it; afterward, they can move the slider position freely to decide on their investment. The variable payment is based on this investment task; that is, we incentivize this task separately.\textsuperscript{12} Participants know that the computer will randomly select 1 of their 10 investment decisions after having completed the final questionnaire, and the performance of that chosen investment decision determines their completion bonus payment. The participants final wealth is the value of the selected stock investment after 1 year (beyond the ending time of the price chart) plus the money in the bank account. The return of the selected stock investment after 1 year is based on this stock’s actual return in the subsequent year, that is, CRSP data for calendar year 2018 for study 1. The participant received US$0.001 for every monetary unit of final wealth of the selected decision in the investment task.

For the price chart stimuli used in study 1, we select 1,000 price charts based on all 8,453 companies included in the CRSP database for calendar year 2017. We first remove double entries, firms with incomplete price data, firms with negative prices, penny stocks (i.e., prices below US$5), and incomplete stock-year combinations. That filtering leaves 4,246 stocks.

We then form 10 times 10 groups for the 4,246 stocks based on two criteria. The first criterion is the stock’s return within 2017. Each stock is allocated to one decile based on its return in 2017. The second criterion is the degree of convexity and concavity of the price path. Using this measure guarantees a wide variety in price path shapes including convex and concave shapes but also clear trends (positive and negative), and perhaps also associated variation in visual

\textsuperscript{11} For a detailed description of the questions and an example of the interface, we refer readers to Section 4 of the Internet Appendix. In total, each participant was exposed to 10 price charts. The two price paths we do not analyze here are artificial price paths based on geometric Brownian motion (GBM).

\textsuperscript{12} This is one of the provisions we include to account for the recent MTurk data quality debate (Kennedy et al. 2018).
Decision weights for experimental asset prices

salience.\textsuperscript{13} Within each return decile, every stock is assigned to a convexity score decile. That procedure creates 100 groups in 10 return deciles crossed with 10 convexity deciles. We then randomly pick 10 stocks from each of the 100 groups, an approach that creates a set of 1,000 sampled stocks. This procedure guarantees that selected stocks are “representative” of the entire sample along the two dimensions, return and convexity.\textsuperscript{14}

Each participant sees eight of those 1,000 different stock charts in randomized order. This generates a data set of 4,000 choices. In our randomization procedure, we make sure that each of the 1,000 charts is seen by exactly four participants.

In the final questionnaire, we include demographics, questions on risk taking, and general traits of decision-making. Moreover, we control for attention and understanding by including a set of quiz questions after the experiment. Internet Appendix Section 4 offers the details.

As Internet Appendix Section 4 also shows, the recruited MTurk sample overall matches the U.S. population exhibiting a somewhat overrepresentation of high-education groups and younger individuals. This is consistent with previous literature documenting that MTurkers are quite representative of the population of U.S. internet users (Ross et al. 2010; Dellavigna and Pope 2017).

3.2 Laboratory study

Study 2 is designed to test whether visual salience affects decision-making even under highly controlled conditions where the information provided by the graphical representation of stock returns has no special relevance for a rational decision maker. It is a dynamic laboratory experiment in which artificial stock returns are simple: prices increase or decrease with the same (known) probabilities each period. This return generating process is known to the participants.\textsuperscript{15} Thus, participants have the information relevant for an optimal investment decision, such that they do not need to rely on the visual information from the price paths. As a consequence, visual salience could possibly become irrelevant for predicting investments in this study. The point of this study is to probe the boundaries of visual salience, to determine whether there is a controlled domain in which visual salience is irrelevant. Put in psychological language, this study tests whether the influence of bottom-up visual salience of low-level features is overwhelmed or inhibited by the top-down goals of deciding whether to invest, and memorized knowledge of the return-generating process.

\textsuperscript{13} The results of previous experiments indicate that historical price paths with more pronounced convex or concave shapes attract higher attention of investors (Grosshans and Zeisberger 2018; Nolte and Schneider 2018; Cohn et al. 2015).

\textsuperscript{14} In Internet Appendix Section 4, we show four exemplary price paths chosen based on very low or high convexity score buckets or based on very high or low return buckets.

\textsuperscript{15} These participants are all attending or have attended advanced statistics classes, so they are familiar with concepts like expectation and variance.
The experiment for study 2 is computer based and has three parts: comprehensive instructions, the main experimental task, and a concluding questionnaire (see Internet Appendix Section 6).

In the main part, participants repeatedly decide how much of their endowment to invest in a risky asset (the part of the endowment that is not invested earns zero return). As described above, the risky asset is a binary lottery that has one positive and one negative outcome and outcomes are independent across time periods. For a detailed specification of the set of binary lotteries, we refer readers to the instructions of the experiment in Section 6 of the Internet Appendix. Expected return, standard deviation, and skewness of outcomes is constant for a given asset, but varies between assets.

Each participant sees 18 different asset price series in randomized order. For 15 of the 18 assets, the participants first observe 15 periods in which a price path develops, one period at a time, based on the outcomes of the binary lottery. In these first 15 periods, they cannot make any investment decisions. Over the subsequent 15 periods (periods 16–30), they make a sequence of 15 investment decisions after each period. Immediately after they confirmed their decision in period \( t \), the return in period \( t \) is realized and displayed, and they make another decision for period \( t+1 \). 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 two-thirds of the 15 assets. 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 have a greater visual variety in price path shapes, although the asset returns only have two outcomes. The remaining one-third are non-pre-constructed charts with 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 are paid according to a physically random realization of 1 of the 18 decision rounds (participants draw numbered ping-pong balls). The experiment concludes with a questionnaire on demographics and self-perceived behavior.

We recruited 275 undergraduate and graduate business students at a university in Germany. The experiment was conducted under controlled conditions.
Decision weights for experimental asset prices

conditions in the computer labs of the university in 18 separate sessions. On average, participants took 61 minutes to complete the entire experiment (including instructions and questionnaire), and average earnings were €16.60 (about US$18.68 at the time of the experiment). For comparison, a student research assistant earned €10.00 per hour at that time.

3.3 Empirical strategy
To set up the full model in Equation (1), we still need to define the value function. We rely on the CPT value function, which incorporates three different psychologically motivated aspects: reference dependence, loss aversion, and diminishing sensitivity. The functional form we use is

\[ v(x_k) = \begin{cases} 
  x_k^\alpha & \text{for } x_k \geq 0 \\
  \lambda (-x_k)^\alpha & \text{for } x_k < 0,
\end{cases} \tag{3} \]

with \( \lambda > 1 \) and \( \alpha < 1 \) (Tversky and Kahneman 1992).

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

The parameter \( \alpha \) measures the diminishing marginal sensitivity that the decision maker has toward larger gains or losses. We start with the common choice of \( \alpha = 0.88 \) and \( \lambda = 2.25 \) and test the robustness of our results to different values in Internet Appendix Section 7.

Both the decision weights and the value function affect the evaluation of the historical price path. Since the visual salience measure only enters the functional through the decision weights, we propose an additional measure based on correlations of decision weights and returns. This correlation measures the direct influence of decision weights on a price path’s attractiveness independent of the features of the value function, like curvature or loss aversion.\(^8\) This is like a “nonparametric” measure because it is independent of the value function parameters \( \lambda \) and \( \alpha \). Note that we can rewrite Equation (1) as

\[ V(X) \propto E(\pi_k v(x_k)) \]

\[ V(X) \propto E(\pi_k) E(v(x_k)) + Cov(v(x_k), \pi_k). \]

To capture the direct influence of decision weights on \( V(X) \), we isolate the covariance term. The correlation measure between returns and decision weight \( Corr(x_k, \pi_k) \) is just the covariance term normalized by estimated standard deviations.\(^9\) We will use \( Corr \) in the following analyses. The intuition behind

---

\(^8\) When the number of decision weights is much larger than the number of possible returns, we artificially depress any covariational/correlational structure. In study 2, we explore risky assets with only two possible returns. For a meaningful setup in this case, weights collapse to the number of outcomes. This does not hold for the visual salience weights. Thus, we will not use the correlational measure for the analyses in study 2.

\(^9\) For linear utility without loss aversion, the two measures have similar effects on \( V(X) \).
Table 3

Regressions for study 1

<table>
<thead>
<tr>
<th>Model</th>
<th>IA [%]</th>
<th>IA [%]</th>
<th>IA [%]</th>
<th>IA [%]</th>
<th>IA [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>0.0347</td>
<td>0.0456</td>
<td>0.0445</td>
<td>0.0510</td>
<td>0.178***</td>
</tr>
<tr>
<td>(2)</td>
<td>(0.0138)</td>
<td>(0.0124)</td>
<td>(0.0124)</td>
<td>(0.0139)</td>
<td>(0.0423)</td>
</tr>
<tr>
<td>(3)</td>
<td>0.0891</td>
<td>0.0827</td>
<td>(0.0649)</td>
<td>(0.0752)</td>
<td></td>
</tr>
<tr>
<td>(4 - Tobit)</td>
<td>−0.0452</td>
<td>−0.0408</td>
<td>(0.0409)</td>
<td>(0.0468)</td>
<td></td>
</tr>
<tr>
<td>(5)</td>
<td>0.307**</td>
<td>0.307**</td>
<td>(0.139)</td>
<td>(0.0356)</td>
<td></td>
</tr>
</tbody>
</table>

Controls

<table>
<thead>
<tr>
<th>Observations</th>
<th>NO</th>
<th>YES</th>
<th>YES</th>
<th>YES</th>
<th>YES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls</td>
<td>4,000</td>
<td>4,000</td>
<td>4,000</td>
<td>4,000</td>
<td>4,000</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.002</td>
<td>.238</td>
<td>.239</td>
<td>.240</td>
<td></td>
</tr>
</tbody>
</table>

The table reports results from a regression of invested amounts on the correlation of returns (Models (1)-(4)) with visual salience decision weights, CPT decision weights, and salience decision weights. Model (5) reports results for the full setup, but with values instead of correlations. Control variables are the full set of statistical measures discussed in Section 2.3, calculated on path level. In Model (4), we use a tobit regression to account for accumulation of participants who did not invest in that stock. In all regressions, we use fixed effects to control for participant-specific effects. All variables are standardized for easy comparison of coefficient sizes and to enhance readability of the table. Robust standard errors clustered by participant level are reported in parentheses.* $p < .1$; ** $p < .05$; *** $p < .01$.

4. Results

4.1 Results: Study 1

Study 1 intends to test whether the visual salience model can predict investment choices of experimental participants, based on empirical price charts. Table 3 presents the regression results of valuations on invested amounts for all 1,000 CRSP-based price charts.

Model (1) of Table 3 includes the correlation measure for the visual salience decision weights, to remove the possible effect of different value functions. We see that the correlation created by stand-alone visual salience is highly predictive for invested amounts. A suggestion from our statistical
feature analyses in Section 2.3 is to test whether the predictability of visual salience remains after controlling for statistical features of the price paths. Model (2) presents the results, showing that visual salience still significantly predicts invested amounts. We also want to address whether visual salience predictability goes beyond standard models like CPT (Tversky and Kahneman 1992) as applied by Barberis, Mukherjee, and Wang (2016) or salience theory (Bordalo, Gennaioli, and Shleifer 2012; 2013) as applied by Cosemans and Frehen (2021). Therefore, we run regressions for the correlational measure where we add the CPT-score and salience-score for the different price paths as additional controls (Model (3)). The results confirm a significant and positive relation between visual salience and the invested amount. Model (4) is a tobit regression to account for the clustering of invested amounts at zero; this has little impact on the visual salience coefficient. One possible specification of the value function is a CPT-value function. Model (5) repeats the analysis of Model (3) for this specification. The general result remains unchanged (the $R^2$ is the same to two decimal places).

To identify potential drivers of our results, we investigate the impact of visual salience for gains and losses, separately. Interestingly, the observed effect solely stems from losses, that is, only differences in visual salience weights associated with losses significantly affect invested amounts. This distinction holds for all Models (1)–(5) considered in Table 3. This result supports the notion that losses serve as modulators of attention (Yechiam and Hochman 2013). The complete regression table is presented in our Internet Appendix Section 5, where we also provide 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.

4.2 Results: Study 2
Study 2 aims to answer whether visual salience decision weights remain a significant predictor of investments even in a simple, highly controlled environment with full information about the return-generating process. To analyze our data, we use all 15 investment decisions of each decision round (rounds 16–30) as our dependent variable. Since our asset only has binary

---

20 In our MTurk sample, in 13% of all decisions, nothing was invested, and 0.2% of the participants never invested.

21 Here, we present results for the parameterization used in the original paper by Tversky and Kahneman (1992), namely, $\alpha=0.88$ and $\lambda=2.25$. In Internet Appendix Section 7, we show that our results are robust to all $\alpha$-$\lambda$ combinations with $0.7 \leq \alpha \leq 1$ and $0.8 \leq \lambda \leq 3$.

22 Furthermore, visual salience might also influence the formation process of reference points. While this is an interesting direction to dig deeper into the potential of visual salience to learn about the intricacies of financial decisions, it goes beyond the scope of this paper and will be left for future exploration.
Table 4

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5 - Tobit)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IA [%]</td>
<td>IA [%]</td>
<td>IA [%]</td>
<td>IA [%]</td>
<td>IA [%]</td>
</tr>
<tr>
<td>Visual salience</td>
<td>0.0735***</td>
<td>0.0314***</td>
<td>0.0365***</td>
<td>0.0287***</td>
<td>0.0469***</td>
</tr>
<tr>
<td></td>
<td>(0.00927)</td>
<td>(0.00793)</td>
<td>(0.00807)</td>
<td>(0.00838)</td>
<td>(0.0119)</td>
</tr>
<tr>
<td>CPT equal weights</td>
<td>0.125***</td>
<td>-0.0739</td>
<td>-0.0488</td>
<td>-0.0665*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0196)</td>
<td>(0.0524)</td>
<td>(0.0589)</td>
<td>(0.0403)</td>
<td></td>
</tr>
<tr>
<td>CPT</td>
<td>0.0183</td>
<td>0.0163</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0153)</td>
<td>(0.0140)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salience</td>
<td>0.00524</td>
<td>0.00719</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00801)</td>
<td>(0.00794)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.0819***</td>
<td>0.128***</td>
<td>0.233**</td>
<td>0.172</td>
<td>0.469***</td>
</tr>
<tr>
<td></td>
<td>(0.00859)</td>
<td>(0.0122)</td>
<td>(0.0992)</td>
<td>(0.119)</td>
<td>(0.114)</td>
</tr>
</tbody>
</table>

The table reports regressions of invested amounts on the visual salience model. In Models (2)-(5), we also include a CPT value function with equal (1/n) decision weights to control for the effect of value function specification and isolate the decision weights’ contribution. Control variables are expected return of the binary lottery, its standard deviation, its skewness, and the average realized return. In Model (4), we add values derived from CPT (Barberis, Mukherjee, and Wang 2016) and salience theory (Cosemans and Frehen 2021) as controls. Model (5) is 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 by participant level are reported in parentheses. *p < .1; **p < .05; ***p < .01.

outcomes, analyzing the correlation between weights and returns as in study 1 is not valid. So no parallel analysis with weight-returns correlations (as in Table 3) will be reported. Instead, we focus on the CPT value function.

We run regressions with and without control variables. In all our regressions, we include the average realized return for each of the 15 periods to account for level effects, which could result in, for example, the disposition effect. As additional control variables, we include expected return, standard deviation, and skewness of the different price paths. These controls are similar to the statistical features we used as control variables in Study I.23

Table 4 reports regression results for the visual salience model (Model (1)), which shows that visual salience coefficient is positive and highly statistically significant. Clearly, our results depend on the specification of the value function. To account for this, in a first step we include a CPT value function with equal decision weights as a control variable in Model (2). Thereby, we create a baseline effect that can be attributed to the particular specification of the value function, isolating the effect of visual salience decision weights. The visual salience score retains its significance despite this. In Model (3), we add controls and the picture for visual salience does not change, but CPT with equal weights loses its significance, which is plausible since the value function is largely

23 Since we only have binary outcomes and therefore a limited set of possible features in study 2, we cannot calculate all features used in study 1.
Decision weights for experimental asset prices

based on risk and return characteristics. In Model (4), we also include values based on CPT and salience theory decision weights in the spirit of Barberis, Mukherjee, and Wang (2016) and Cosemans and Frehen (2021). We see the same pattern we have observed in study 1 with respect to the visual salience measure. Its coefficient is highly significant and has positive explanatory power for investments. In Model (5), 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.

In Section 7 of our Internet Appendix, we test the sensitivity of our analysis to the values of the curvature ($\alpha$) and loss aversion ($\lambda$) parameters in the value function. Curvature has hardly any impact for the influence of visual salience on investment decisions. The effect of visual salience is also relatively independent of the chosen loss aversion parameter. Only for parameter values indicating loss seeking or no loss aversion ($\lambda \leq 1$) we observe an impact on the performance of visual salience.

5. Discussion

5.1 Bottom-up and top-down decision weights

Bottom-up visual salience can predict invested amounts beyond statistical features and established models like CPT and salience theory. Attention is guided by both universal and rapid bottom-up processes, and more deliberate top-down considerations. Most (normative or descriptive) models of financial decision-making have an implicit assumption of deliberate, top-down guided agents. These agents apply what they know (models, goals, experience) and what they expect to perceive. They also fill in the blanks, calculating statistical properties like mean, standard deviation, or skewness of return distributions, which are prominent in models of financial decision-making. We suggest that models like CPT in the sense of Barberis, Mukherjee, and Wang (2016) or salience theory as applied by Cosemans and Frehen (2021) capture perceptual and behavioral effects of this top-down information processing, while visual salience captures the bottom-up part of the decision-making process.

Our analyses of visual decision weights and statistical features in Section 2.3 suggest that (bottom-up) visual attention cannot be exclusively explained by (top-down) statistical features. We find further support for this by regressing our correlational measure—which is the part of visual salience valuation that can be attributed directly to the decision weights—on the three most common statistical properties of (past) return distributions, namely, mean return, standard deviation, and skewness of returns. The resultant $R^2$ is exceptionally low at 0.28%. The explanatory power is much higher when running the same analysis for correlational measures of CPT ($R^2 = 87.98\%$) and salience theory ($R^2 = 66.71\%$). This supports our assertion that these theories are closely

\footnote{Internet Appendix Section 9 presents the regression tables.}
Figure 3
Optimizing bottom-up and top-down combinations
This figure shows the $R^2$ of regressions predicting investments for combinations of visual salience and CPT (left graph) and salience (right graph) decision weights. The highest $R^2$ are attained at a weight of 60% and 97% on visual salience, respectively.

linked to top-down decision-making captured by statistical properties of returns. Moreover, visual salience decision weights are practically uncorrelated with CPT ($\rho = -0.0086$) or salience theory ($\rho = -0.0034$), while the weights of these two theories are highly correlated with one another ($\rho = 0.8032$).

If—as these analyses suggest—visual salience captures bottom-up decision weights, and the other two are more related to top-down valuation processes, combining the two could create a good synthesis. The simplest approach would be a linear combination of the decision weights. We optimize across all possible linear combinations to find the highest attainable $R^2$ predicting invested amounts.\(^{25}\) For visual salience combined with CPT, the optimized resultant $R^2$ is 7.43%, with 60% of the weight on visual salience and 40% on CPT decision weights (Figure 3). For visual salience combined with salience theory, the highest attainable $R^2$ is even higher at 7.91%, where most of the weight (97%) is on visual salience decision weights. This analysis suggests that a combination of different approaches can indeed improve explanatory power, as $R^2$ of standalone models are 7.04% (visual salience), 7.03% (CPT), and 2.9% (salience theory).\(^{26}\)

5.2 Visual salience and belief formation
In our model, we argue that bottom-up visual salience enters decision-making by distorting the probabilities after having transformed the outcomes from past dates into possible future states. However, we have not discussed whether the predictability observed so far might be driven in the preceding step, by changes in beliefs about the future return distribution. Beliefs about the future

\(^{25}\) We do not include control variables in this setup since we are only interested in the explanatory power of each setup. Internet Appendix Section 9 presents the full regression tables.

\(^{26}\) $R^2$ for salience is comparably low, which might be because it was designed for context-dependent decision-making. It is well-established in the literature that—in different setups, for example, where investors choose among several different investment options—it outperforms many other decision models.
return distribution are usually modeled relying on past returns (for a survey, see Assenza et al. 2014). If visual salience affects this transformation of past returns into beliefs about the future return distribution, then the more visually salient states are indeed considered to be more likely, and visual salience enters the valuation through the expectation and belief channel. Many recent studies have tried to understand belief formation regarding future outcomes (e.g., Greenwood and Shleifer 2014; Barberis et al. 2015; Cassella and Gulen 2018; Da, Huang, and Jin 2021). Key findings are extrapolation and overreaction, for example, an overestimation of recent trends and shocks, stickiness of expectation adjustment and underreaction to information (see, e.g., Afrouzi et al. 2019).

We conduct two analyses to test the hypothesis that visual salience affects the investment choices of our experimental participants by changing their beliefs about the future return distribution. To quantify whether beliefs are affected by visual salience, we employ two standard metrics: expected return and risk (Bloomfield and Hales 2002). Our main experimental findings about investment behavior are based on a cross-sectional analysis. To be consistent, we stick to this perspective and asked participants in study 1 to also state return expectations for the following 12 months on a scale between $-100\%$ and $+100\%$, as well as how risky they deem an investment into that stock on a scale from 1 (not risky) to 10 (very risky). In a first analysis, we test whether visual salience is able to predict expected return and risk. Second, we test whether visual salience decision weights capture more than beliefs when predicting investment choice, by explicitly controlling for self-stated beliefs about future risk and return.

Inspired by models on extrapolation bias (Cassella and Gulen 2018; Afrouzi et al. 2019), we approximate the return expectation as a weighted-average of the observed experienced past data for each stock $i$, that is, we regress the elicited expectation of each participant $j(\text{ExpRet}_{i,j})$ on the weighted average of past returns. The weighting is based on visual salience and a parameter $\gamma$. The parameter $\gamma$ modulates the degree of responsiveness to visual salience; that is, we search for the $\gamma$ that maximizes the $R^2$ of the following regression over each price path $i=1,\ldots,1,000$ and all participants $j=1,\ldots,500$:

$$\text{ExpRet}_{i,j} = b_0 + b_1 \mu_{i}^{VS} + \text{controls} + \varepsilon_{i,j}$$

with

$$\mu_{i}^{VS} = \frac{1}{\sum_{k=1}^{250} (\pi_{k,i}^{VS})^{\gamma}} \sum_{k=1}^{250} (\pi_{k,i}^{VS})^{\gamma} x_{k,i},$$

where $\gamma = 0$ indicates that the participants do not respond to visual salience at all when forming their return expectation, $\gamma \in (0, 1)$ a dampened response to visual salience.

27 Testing a longitudinal relation is left for future research.
Figure 4
Predicting return expectations and risk with visual salience
This figure shows the R² (y-axes) of regressions predicting expected return (left graph) and risk (right graph) as a function of the exponentiated weight on salience, γ (x-axes). The highest R² is at γ=0.27 for predicting return expectations and at γ=0.64 for risk expectations.

Figure 4 presents the results for expected return and risk regressions varying γ, the weight on visual salience, from (0,1). The left panel depicts the R² on the y-axis for expected returns. The model that best fits the elicited return expectations is reached for γ=0.27. This implies that a model that assumes fairly low responsiveness to visual salience when forming their return expectation best fits the elicited expectations data. This low responsiveness is sensible, since expectations are likely based to a large extent on path

28 We do not control for actual realizations of a stock’s return distribution since doing so would create high levels of multicollinearity with μVS and σVS for low levels of γ.

29 We ran regressions up to γ = 10 to check for additional local optima but found none.
Decision weights for experimental asset prices

characteristics like the average return in combination with recent trends Cassella and Gulen (2018). Average return is a top-down perception that, by definition, erases the differential salience of each individual return in the time series. The results suggest that top-down perception largely overrides bottom-up initial visual salience-guided attention.

In contrast, when turning to risk expectations we see that the best performing model is reached for $\gamma = 0.64$. That is, a model that assumes that experimental participants are moderately responding to visual salience weights explains risk beliefs most accurately. Bottom-up visual salience can be sensitive to contrast and line orientation novelty. This is sensible if the ups and downs of a price path are more visually salient, and at the same time clearly influence risk perception.30

The final question we need to answer is whether visual salience is only a good proxy for risk perception and does not enter the decision-making process through another channel, namely, via the channel assumed in this paper: a distortion of decision weights. To test this, we add expected return and risk as additional control variables to our main regression analyses (from previous Table 3) in Table 5. Coefficients for expected return (positive) and risk (negative) are plausible and significant, indicating that we indeed capture relevant beliefs with these variables. Predictability of our visual salience model retains its significance with a similar coefficient, indicating that it indeed captures more than just beliefs (of the first two moments of return distributions).

Much is certainly left to be explored about the relation between visual salience and belief formation. This includes, but is not limited to, the channel through which visual salience affects risk to a larger degree than expected return beliefs. This will help researchers better understand how bottom-up visual salience affects decision-making, but will be left to future research.

6. Conclusion

Attention toward salient features is reliably predicted by their bottom-up properties of stimuli, independent of goals. We use a particular algorithm to predict visual salience of prices in a sequential price path. 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 overweight 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, Lang, and Yetman 2009), or during a particularly long streak (Raghubir and Das 2010).

30 Another feature that attracts attention is “Bayesian surprise,” namely, the sudden appearance of a feature that was previously rare in a sequence of images (Itti and Baldi 2009; Baldi and Itti 2010). This could play a role in study 2, where the sequence of prices are presented over time, but our design is not well-equipped to analyze Bayesian surprise carefully (since it requires a clear prior belief about what features are expected).
Table 5
Regressions for study 1

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IA[%]</td>
<td>IA[%]</td>
<td>IA[%]</td>
</tr>
<tr>
<td>Corr(Return, VS weights)</td>
<td>0.0378***</td>
<td>0.0386***</td>
<td></td>
</tr>
<tr>
<td>VS (value)</td>
<td>(0.00989)</td>
<td>(0.01000)</td>
<td>(0.129***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0344)</td>
</tr>
<tr>
<td>Expectations: return</td>
<td>0.376***</td>
<td>0.337***</td>
<td>0.336***</td>
</tr>
<tr>
<td></td>
<td>(0.0225)</td>
<td>(0.0230)</td>
<td>(0.0229)</td>
</tr>
<tr>
<td>Expectations: riskiness</td>
<td>−0.315***</td>
<td>−0.275***</td>
<td>−0.274***</td>
</tr>
<tr>
<td></td>
<td>(0.0196)</td>
<td>(0.0201)</td>
<td>(0.0201)</td>
</tr>
<tr>
<td>Corr(Return, CPT weights)</td>
<td>0.0984*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0534)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corr(Return, salience weights)</td>
<td>−0.0535</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0334)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPT (value)</td>
<td></td>
<td></td>
<td>0.0923*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0533)</td>
</tr>
<tr>
<td>Salience (value)</td>
<td></td>
<td>−0.0525</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0331)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.533***</td>
<td>0.460***</td>
<td>0.459***</td>
</tr>
<tr>
<td></td>
<td>(0.0500)</td>
<td>(0.0502)</td>
<td>(0.0503)</td>
</tr>
<tr>
<td>Controls</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>4,000</td>
<td>4,000</td>
<td>4,000</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.463</td>
<td>0.484</td>
<td>0.484</td>
</tr>
</tbody>
</table>

The table reports results from a regression of invested amounts on the correlation of returns (Models (1) and (2)) with visual salience decision weights, controlling for return and risk expectations. Model (3) reports results for the full setup, but with values instead of correlations. Control variables are the full set of statistical measures discussed in Section 2.3, calculated on path level. In all regressions, we use fixed effects to control for participant-specific effects. All variables are standardized for easy comparison of coefficient sizes and to enhance readability of the table. Robust standard errors clustered by participant level are reported in parentheses. * $p < .1$; ** $p < .05$; *** $p < .01$.

We repurpose machine learning techniques, adapted to predict what people look at in any kind of image, to predict salient points in price paths. To the extent that human perception of visual stimuli is universal and consistent across many domains, an all-purpose algorithm might also predict the salient aspects of financial stock price paths. However, visual salience tackles the first few seconds of “bottom-up” (goal-free) perception only. This does not mean it will predict salience over longer periods of deliberation, or that bottom-up salience predicts decisions with a longer decision-making process where attention is guided by specific goals.

We first confirmed that SAM predictions of visual salience of price paths correlate closely with new eye-tracking data. Next, we relate SAM predictions with statistical features of the historical return distribution. This also helps to better understand the “black box” SAM and therefore is related to explainable AI. In particular, we use feature relevance to enhance explainability and show that only 20% of the variance in SAM’s implied decision weights can be explained by these statistical features.

We then found in two experimental studies that the visual salience model has predictive power to explain actual investment decisions. This predictive power is evident for real stock price charts, and for the easiest possible depiction of artificial price charts with repeated binary lotteries. The visual salience model
Decision weights for experimental asset prices

retains its explanatory power controlling for a wide variety of statistical features of the price charts, and for other economic models that focus on different aspects of attention and salience like perception of probabilities and return differences.

Of course, in natural settings investors likely look at many different pieces of information besides price charts. Our experiments present a stylized setup where the price path is the only information source to isolate the effect of bottom-up visual salience. While bottom-up processing assembles and integrates the sensory information presented in a price path, top-down processing interprets this information based on models, ideas, and previous experience. In a further set of analyses, we show that combining the visual salience model with models like CPT (Tversky and Kahneman 1992) in the sense of Barberis, Mukherjee, and Wang (2016) or salience theory (Bordalo, Gennaioli, and Shleifer 2012) as applied by Cosemans and Frehen (2021) improves predictive power. These latter theories capture the perceptual and behavioral effects of top-down information processing.

Visual salience is grounded in basic human visual perception and thus is directly measurable via eye-tracking and predictable via algorithms like SAM. This allows us to analyze the impact of visual salience in any economic domain in which visual features of the provided information and alternatives may play a role (e.g., Li and Camerer 2021). 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 ignoring certain features. One further layer is belief formation. Beliefs about the future return distribution are usually modeled relying on past returns (for a survey, see Assenza et al. 2014). If visual salience affects this transformation of past returns into beliefs about the future return distribution, then the more visually salient states are indeed considered to be more likely, and visual salience enters the valuation through the expectation and belief channel. We show in a static environment that the influence is low for predicting return expectations, and moderate for risk expectations. Analyzing belief formation of the future return distribution in a dynamic context is left for future research.

In general, algorithms like SAM can give portable, parameter-free answers to the question, “What do most people notice first?” When nothing visual is displayed, the question and answer are uninteresting. However, this question could be interesting in many other cases: the presence of display effects in online

31 Camerer and Loewenstein (2003) called this “generality” (following Stigler 1965), and several years later Rabin (2013) coined the similar term “portability.”
shopping; differences in visual financial disclosure; detailed documents like 10-K filings; pictures in press releases and annual reports (e.g., does visual salience of a CEO in the report predict hubris or weak governance?); differences in the length of history for which returns are displayed (and formatted, for example, the y-axis could make a difference); and so on. Surely some of these questions are interesting for financial economics; and happily, computer vision scientists have already done one of the most difficult parts—predicting salience—for us.

References


Decision weights for experimental asset prices


Decision weights for experimental asset prices


