# Testing for salience effects in choices under risk*

Carsten S. Nielsen†‡ Alexander C. Sebald† Peter N. Sørensen†

January 22, 2018

DRAFT

**Abstract**

Recent economic literature acknowledges the fact that our attention is limited and predictively drawn to specific features of our environment. We experimentally investigate a prominent approach to model limited attention in decision making: Salience Theory, proposed by Bordalo et al. (2012b, 2013a). We specifically design a risky-choice experimental environment which allows us to control for a large array of alternative theories of behavior. Our experimental results are in line with the behavioral hypothesis of Salience Theory which posits that decision makers base their decisions on decision weights that distort the prospects' actual probabilities in favor of the salient consequences.

**Keywords:** Salience, attention, risky behavior

**JEL-Classifications:** D81, D91, G11, G41

*We are grateful for comments from seminar participants at Bocconi, University of Cologne, ECARES (Bruxelles), University of St. Gallen, University of Copenhagen, Lund University, Aarhus University, University of Southern Denmark (Odense), Université Laval (Québec City), NYU (Abu Dhabi). Furthermore we thankfully acknowledge the financial support from the Danish Council for Independent Research in Social Sciences (Grant ID: DFF-4003-00032).

†Department of Economics, University of Copenhagen, Øster Farimagsgade 5, Building 26, DK-1353, Copenhagen K, Denmark.

‡Corresponding author: alexander.sebald@econ.ku.dk

# 1 Introduction

Classical economic theory builds on a simple but powerful model of rational behavior: People are assumed to have unlimited cognition and make choices so as to maximize their wellbeing by attending to all available information. However, since the amount of information we are continuously confronted with is vast and cognitive resources are limited, our attention and perception is restricted, implying that we often neglect important aspects of the environments in which we act. The fact that our attention is limited and predictively drawn to specific features of our environment has recently stirred great interest in economics [see e.g. Kőszegi and Szeidl (2012), Bordalo et al. (2012b), Woodford (2012), Schwartzstein (2014), Dertwinkel-Kalt et al. (2017), Bordalo et al. (2017) etc].

One prominent theory in this strand of the literature is Salience Theory. Starting from basic insights on human perception, Pedro Bordalo, Nicola Gennaioli and Andrei Shleifer (Bordalo et al., 2012b,a, 2013b,a, 2015, 2016) have recently developed 'Salience Theory' to explore economic consequences of the way salience captures attention. In their theory, the salience of choice features systematically impacts on choices via a perturbation of individual preferences. Their core intuition rests on two psychophysical concepts of perception, which in combination have become known as the Weber-Fechner law (Dzhafarov, 2001; Dzhafarov and Colonius, 2011).

In our analysis we construct an experiment to isolate and test the choice effect predicted by Salience Theory. Individuals are asked to solve simple choice problems of allocating wealth to a risky investment. We rule out systematic effects by a majority of theories of choice under risk by creating choice problems that have economically equivalent consequences. We are still able to vary the salience of the risky choice, and hence able to measure an effect predicted by Salience Theory.

The first part of the the Weber-Fechner law is *Weber's principle*. It states that dissimilarity between two stimuli magnitudes is determined by the ratio of the large magnitude to the low. For example, the difference between 11 and 10 is perceived similar to a difference between 22 and 20. The second part is *Fechnerian sensitivity*. This refers to the implication of Weber's principle that there is diminishing sensitivity to a given difference in stimuli magnitude. It is easier to discriminate between a difference between 11 and 10, than a difference between 16 and 15.

Let us illustrate the main idea of our experiment by explaining two treatments, A and B. In treatments A and B subjects are respectively endowed with \$16 and \$11. They choose to bet an amount between \$0 and \$10 on a risky lottery. There is a 60% chance to receive \$1.1 per dollar bet in treatment A and \$1.6 in treatment B. They have a remaining 40% chance to receive \$0.1 per dollar bet in A and \$0.6 in B. For each dollar not bet, out of the \$10, subjects receive the risk-free \$0.4 in A and \$0.9 in treatment B. Finally, the part of the endowment that cannot be bet (i.e. \$6 in treatment A and \$1 in B) keeps its value.

This setup is deliberately created such that the state-dependent wealth consequences of any amount $\alpha$ bet on the risk lottery are the same for the two treatments. If subjects decide to bet amount $\alpha$ on the risky lottery, then in both treatments the payoff is $\$10 + \$0.7\alpha$ with probability 0.6 and $\$10 - \$0.3\alpha$ with probability 0.4.[1] In other words, the prospects are exactly the same given any bet, and the choice sets of bets from \$0 to \$10 are identical.

Salience Theory posits that the Weber-Fechner law predicts which consequence is most salient, and that the relative salience of consequences influences choices. In their theory for risky decision making (Bordalo et al., 2012b, 2013a), consequences are defined by payoffs. The relative payoffs to the risky bet differ in our treatments. The bad consequence is most salient in treatment A since $\$1.1/\$0.4 < \$0.4/\$0.1$, while the good consequence is most salient in treatment B since $\$1.6/\$0.9 > \$0.9/\$0.6$. More salient consequences get assigned larger decision weights. In treatment B, the good consequence is more salient and weighted, causing betting to seem more attractive.

Salience Theory thus predicts the effect that subjects will bet more in treatment B. Recalling that prospects are constant across the two treatments, other theories of decision making predict zero effect of the treatment. This is the basic identification strategy of our experiment.

In our experiment we invited subjects to a sequence of four treatments structurally similar to the illustration above. The four treatments vary the payoff ratios and hence the extent of salience, from treatment level '0' when the bad consequence of the risky lottery was most salient to treatment level '3' when the good consequence of the risky lottery was most salient. At some point between treatment levels 1 and 2, the two consequences are equally salient, in theory. In our setup, participants are serially assigned to one of

[1] In Treatment A, the attained wealth will be $\$6 + 1.1\alpha + \$0.4(\$10 - \alpha)$ if the good consequence obtains and $\$6 + \$0.1\alpha + \$0.4(\$10 - \alpha)$ if the bad consequence obtains. In Treatment B it will be $\$1 + \$1.6\alpha + \$0.9(\$10 - \alpha)$ and $\$1 + \$0.6\alpha + \$0.9(\$10 - \alpha)$, respectively.

the possible 24 orders in which the four treatments can be presented.

If our design is successful in manipulating the Weber-Fechner law, Salience Theory predicts that subjects will bet more on the risky lottery at increasing treatment levels. Our experimental results show that the observed behavior strongly supports this prediction. On average participants in our experiment bet about 77% of the amount that they can bet on the risky lottery at the baseline level '0' and increase their bet towards treatment level '3' by on average 18%. To go beyond these experimental results, we also use a structural model to estimate a 'deep' salience parameter. Estimating such a parameter can aid in quantifying the impact of salience on outcomes beyond those implied by our experimental design. We find that increasing the payoff contrast by one percent will increase the perceived ratio of lottery probabilities by 0.39%.

As mentioned before, as our choice problems have identical economic consequences, our experimental set-up controls for a large set of theories which could otherwise be associated with betting behavior in general. In particular, it controls for any decision theory that takes prospects as input. This includes, for example, Expected Utility Theory (Von Neumann and Morgenstern, 1944; Savage, 1954), Regret Theory (Bell, 1982; Loomes and Sugden, 1982), Disappointment Theory (Bell, 1985; Loomes and Sugden, 1986; Gul, 1991), and Similarity-Based Theories (Rubinstein, 1988; Leland, 1994). It is also impossible for common consequence and common ratio effects (Allais, 1953) to explain any effects in our experiment. The common consequence effect would need a third consequence, while the common ratio effect would need to change probabilities by proportion. The two-consequence feature of our experimental design also controls for more general theories that loosen (Chew, 1983; Fishburn, 1983), or even discard (Machina, 1982), the 'independence axiom' of Expected Utility Theory.

Another obvious candidate theory is of course Prospect Theory. As in Prospect Theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), Salience Theory à la Bordalo et al. (2012b, 2013a) implies a form of 'narrow framing,' which occurs when people evaluate a risk separably from other concurrent risks. This manifests itself in that probabilities (via the Weber-Fechner law) are biased by the payoffs, whereas in Prospect Theory value is derived from 'gains' and 'losses' measured relative to some reference wealth level. Although these are very different concepts, both could potentially explain the illustrated change in risk attitudes across Treatment A and Treatment B in our set-up. For Prospect Theory it is unfortunately often unclear how to define precisely what a gain or loss is, not least because Kahneman and Tversky offered relatively little guidance on how the reference wealth level is determined. However, it turns out that our experimental design controls for two significant attempts to pin this down.

First, Kőszegi and Rabin (2007) propose that people have rational expectations about potential wealth consequences, and derive utilities from the difference between actual consequences and these expectations. Rationality implies that subjects' expectations will match the distribution of consequences they face prior to deciding on the amount to bet. In the context of our example this means that the reference wealth level will be given by the identical $0.6(\$10 + \$0.7\alpha) + 0.4(\$10 - \$0.3\alpha)$ for both treatments. A rational expectation reference point is therefore also unable to explain a change in behavior. Second, Barberis and Huang (2008) define the reference wealth level as the 'status quo' to account for the observed skewness in the distribution of asset returns. Defined like this, the status quo refers to the situation in which subjects bet nothing in our task. If subjects in our example bet nothing in treatments A and B, then in addition to the part of the endowment that could not be bet, i.e. \$6 in A and \$1 in B, they receive the risk-free earnings \$4 and \$9 respectively. The status quo reference wealth level is thus \$10 in both treatments, and therefore cannot explain a treatment effect on betting behavior.

There remains a third adaptation of prospect theory which could predict an effect in our design. In their effort to explain the equity premium puzzle, Benartzi and Thaler (1995) propose that investors in asset markets judge their portfolios by the value of their holdings and not their overall wealth levels. In analogy to Salience Theory, this adaptation of prospect theory implies that subjects allocate more to the risky lottery in treatment B compared to A. In order to assess how well this alternative theory accounts for our experimental data, we estimate a structural model that embeds their idea. It turns out that the obtained estimates are far from earlier studies on this version of prospect theory, making it an unlikely explanation for the observed treatment effect.

Experimental studies on Salience Theory and risk taking are few. Bordalo et al. (2012b) themselves provide preliminary evidence using unincentivized survey experiments conducted using Amazon Mechanical Turk, an online marketplace hosted by Amazon.com. They show that Salience Theory successfully can explain data collected via surveys on Allais common consequence and common ratio effect problems. They also find that Salience Theory can account for survey data on prospects that differ by a mean preserving spread. Their findings on Allais' common consequence problem have since been successfully replicated in an incentivized laboratory experiment by Frydman and

Mormann (2017). Different to our analysis, Frydman and Mormann (2017) use a series of binary choice decisions situations à la Allais and manipulate the correlation between the different lotteries while keeping the marginal distribution of each lottery constant. In their experimental set-up, Salience Theory predicts that the probability of exhibiting the Allais paradox decreases with the correlation between the common consequence lotteries. By keeping the marginal distributions of each lottery constant across the decision situations, their set-up furthermore controls for Expected Utility Theory as well as Cumulative Prospect Theory as alternative explanations for behavior. In two additional experiments, eye-tracking experiments are used to distinguish between Salience Theory and Regret Theory. As the authors state, taken together their experimental analysis lends support for salience driven changes in behavior.

Compared to the evidence presented in Bordalo et al. (2012b) and Frydman and Mormann (2017), our experimental results show that salience effects influence behavior even in situations in which common consequence and common ratio effects are ruled out.

Our analysis is organized as follows. The next section explains and motivates the design of our experiential test. Section 3 presents the experiment and the results, Section 4 discusses the alternative theory due to Benartzi and Thaler (1995), and Section 5 concludes.

# 2 Experimental Design

In this section we formally present the experimental task, characterize salience effects à la Bordalo et al. (2012b, 2013a), and present our identification strategy as well as behavioral hypothesis. Note that we present salience effects both with reference to the rank-dependent as well as the continuous version of Salience Theory.

## 2.1 Task

In each of four decision situations in our experiment, subjects confront a simple portfolio allocation problem. They are given an endowment $e$ and a budget $m$. They are asked which amount $\alpha \leq m$ they want to bet on a risky lottery $L_A$. If $\alpha$ is chosen, subjects face a $p > \frac{1}{2}$ chance of realizing a 'good' outcome with lottery earnings $x_g\alpha$ and a $(1-p)$ chance of realizing a 'bad' outcome with lottery earnings $x_b\alpha$. The amount not bet on $L_A$ is automatically allocated to a risk-free lottery $L_B$ that pays $x_f(m-\alpha)$ no matter the outcome. In all four situations, payoffs have the natural ranking $x_g > x_f > x_b$, and the flat fee $e - m$ keeps its value.

Given this set-up the state-dependent wealth consequence of choosing $\alpha$ in each of the four different decision situations is:

$$c_s = e - m + x_s\alpha + x_f(m - \alpha) > 0. \quad (1)$$

with $s \in \{g, b\}$.

Subjects that do not bet anything earn the amount $x_f m$ with certainty. By betting more, subjects gain more if they are lucky (i.e., the good outcome is realized), but also lose more if they are unlucky (bad outcome). As we will explain in detail below, across decision situations we vary the relative salience of the different outcomes while keeping prospects fixed.[2]

## 2.2 Salience Theory

Salience Theory for risky decision making as characterized by Bordalo et al. (2012b, 2013a) defines 'salience' by a continuous and bounded function $\sigma(x, y)$ on payoffs $x, y > 0$, which satisfies two main properties: ordering and diminishing sensitivity.[3] Ordering says that, if an interval $[x, y]$ is contained in a lager interval $[x', y']$, then $\sigma(x, y) < \sigma(x', y')$. Diminishing sensitivity says that $\sigma(x + \epsilon, y + \epsilon) < \sigma(x, y)$ for all $x, y > 0$ and any $\epsilon > 0$. These two properties are most of the time balanced by assuming that the salience function is homogeneous of degree zero: $\sigma(tx, ty) = \sigma(x, y)$ for all $t > 0$. Such a salience function reflects two stylized facts of human perception. First, our perceptive apparatus is attuned to increases in contrast. Homogeneity of degree zero implies that without loss of generality $\sigma(x, y) = \sigma\left(\frac{y}{x}\right)$.[4] Second, contrast discriminability is easier at lower levels and becomes harder as levels are increased by the same factor. By ordering the salience function $\sigma\left(\frac{y+\epsilon}{x+\epsilon}\right)$ is a decreasing function of $\epsilon > 0$. In particular, diminishing sensitivity is implied because payoff contrast $\frac{y+\epsilon}{x+\epsilon}$ converges towards 1 at a decreasing rate as $\epsilon \to \infty$.

Salience $\sigma$ does not directly effect decision weights, it goes through a filtering mechanism given by the composit attentional weight function $w_s(\sigma)$. When deciding on the

[2]Actual numbers are in Table 1 below.

[3]Note that by adding a symmetry property, namely $\sigma(x, y) = \sigma(-x, -y)$, negative payoffs can also be accommodated. However, this property is irrelevant for our experimental design.

[4]With slight abuse of notation, we follow the convention and denote the restriction of $\sigma(x, y)$ to a one dimensional subset by $\sigma(\frac{y}{x}) = \sigma(\frac{y}{x}, 1)$ for all $x, y > 0$.

amount $\alpha$ to bet in our decision situations, the salient thinker is presumed to weigh the good state by:

$$\pi = \frac{pw_g(\sigma)}{pw_g(\sigma) + (1-p)w_b(\sigma)} \tag{2}$$

and the bad state by $(1-\pi)$.

In the rank-dependent version of Salience Theory the 'attentional weight' put on any consequence $c_s$ is $w_s(\sigma) = \delta^{k_s}$, with $\delta \in (0,1]$ being a parameter which measures the extent to which salience distorts valuations. The outcome state $s$ is said to be more salient than the outcome state $s'$ if $\sigma(x_s, x_f) > \sigma(x_{s'}, x_f)$. Denote by $k_s \in \{1,2\}$ the salience ranking with $k_s = 1$ indicating the more salient state $s$. The salient thinker thus distorts the more and less salient state by $\delta$ and $\delta^2$, respectively. If the two states are equally salient, then they obtain the same ranking. As $\delta \to 0$, the salient thinker will focus all his attention on the more salient state. When $\delta = 1$, there are no salience distortions.

Finally, the salient thinker uses value function $v$ to evaluate consequences, and choose $\alpha$, the amount bet on the risky lottery $L_A$, to maximize

$$V(\alpha) = \pi v(c_g) + (1-\pi)v(c_b). \tag{3}$$

The attitude towards risk as measured by the optimal amount $\alpha^*$ to bet is thus affected by two sources. One is the standard source directly related to the curvature of the value function $v$, the other works indirectly through attentional modulations of the decision weight $\pi$. We will refer to the latter as the 'salience effect'. The salience effect is a form of narrow framing where payoffs, rather than consequences shape the perception of outcome states. This is consistent with, but different from, the framing of prospect theory. In prospect theory narrow framing implies that consequences are perceived as payoff gains and losses relative to a reference point.

Though convenient for theorizing, Bordalo et al. (2012b, 2013a) acknowledge that the assumption of rank-dependent attentional weights is likely too simple. First and foremost, it seems implausible that a slightest deviation from equal salience will change attention by a factor, but stay constant everywhere else. Second, there are some technical issues with rank-dependence that may make it difficult to link rank-dependent salience theory to experimental data (Kontek, 2016).

To ensure that our identification of salience effects is independent of specific functionals, we assume only that attentional weights are positive monotonic transformations of the homogenous salience function. This makes the composite attentional weight function $w_s(\sigma)$ homothetic. Homothetic functions have the ordinal property: $w_s(\sigma(x,y)) \leq w_s(\sigma(x',y'))$ implies $w_s(\sigma(tx,ty)) \leq w(\sigma(tx',ty'))$ for all $t > 0$. For simplicity we omit $\sigma$ when no confusion may arise. Homotheticity thus implies that attentional weights represent salience purely by ranking $w_s(\frac{y}{x}) \leq w_s(\frac{y'}{x'})$. By the ordering property of the salience function, the attentional weight $w_s(\frac{y+\epsilon}{x+\epsilon})$ is a decreasing function of $\epsilon > 0$. This implies that $w_s$ is a convex function: for any positive payoffs $x < y$, and any $\Delta, \epsilon > 0$, the difference $w_s(\frac{y+\Delta}{x+\Delta}) - w_s(\frac{y+\Delta+\epsilon}{x+\Delta+\epsilon})$ is a decreasing function of $\Delta$.[5] By this representation, we have that $w_g(\sigma) = w_g(\frac{x_g}{x_f})$, $w_b(\sigma) = w_b(\frac{x_f}{x_b})$, and the decision weight attached to the good state given by equation 2.

Decision weights sum to one, so the salience value $V(\alpha)$ will satisfy monotonicity, and the maximization problem is straightforward. Given this continuous specification, the salient thinker will bet an amount up to the point where $V'(\alpha) = 0$, at which his preference for risk will be revealed. By equation (3) we therefore have that

$$-\frac{x_f - x_b}{x_g - x_f} = -\frac{\pi}{1-\pi}\frac{v'(c_g)}{v'(c_b)}. \tag{4}$$

The left-hand defines the rate at which attainable wealth can be exchanged between the two outcome states (slope of the budget line), while the right-hand side defines the rate at which a salient thinker is willing to take on risk (slope of the indifference curve). The optimal amount $\alpha^*$ which is bet on the risky lottery ensures equality.

## 2.3 Identification

The identification of the salience effect in our experimental setting is based on a simple idea: We vary payoffs by a constant while keeping the prospects identical across all four treatments. This implies that the salience effect we test for can only be caused by changing sensitivity $\pi$ in (3). In the development of our behavioral hypothesis let payoffs $(x_g, x_b, x_f)$ be expressed as $(\bar{x}_g + \Delta, \bar{x}_b + \Delta, \bar{x}_f + \Delta)$, with $\bar{x} > 0$ being the baseline and $\Delta \geq 0$ a mark-up. We refer to the mark-up $\Delta$ as the treatment variable. Given any bet amount $\alpha$, prospect $(c_g, p; c_b, 1-p)$ is kept invariant in (1) by setting the flat fee at $e - m = m - (\bar{x}_f + \Delta)m$ in our four decision situations. Note that probability $p$ is fixed.

[5] The ordering property implies that $w(\frac{y+\Delta}{x+\Delta}) - w_s(\frac{y+\Delta+\epsilon}{x+\Delta+\epsilon}) > 0$, so the difference is decreasing since $\frac{d}{d\Delta}(\frac{y+\Delta}{x+\Delta}) < \frac{d}{d\Delta}(\frac{y+\Delta+\epsilon}{x+\Delta+\epsilon})$.

Furthermore, to maintain a constant choice set across the four decision situations, we also keep $m$ constant. So, as we vary $\Delta$, it is the endowment $e$ that adjusts to maintain $e - m = m - (\bar{x}_f + \Delta)m$.

By homogeneity of degree zero the salience of the good state is $\sigma(\frac{\bar{x}_g+\Delta}{\bar{x}_f+\Delta})$ while the salience of the bad state is $\sigma(\frac{\bar{x}_f+\Delta}{\bar{x}_b+\Delta})$. We assume that baseline payoffs are such that $\frac{\bar{x}_f}{\bar{x}_b} > \frac{\bar{x}_g}{\bar{x}_f}$ and $\bar{x}_g - \bar{x}_f > \bar{x}_f - \bar{x}_b$, which by the ordering property makes the bad state most salient when $\Delta = 0$. Increasing $\Delta$ will have the payoff ratio in the bad state monotonically decreasing faster towards the asymptote 1 than the payoff ratio in good state. The two will cross at

$$\Delta^0 = \frac{\bar{x}_f^2 - \bar{x}_g\bar{x}_b}{\bar{x}_g + \bar{x}_b - 2\bar{x}_f} > 0$$

where the salience of the two outcome states are the same. When $\Delta > \Delta^0$, the good state is most salient.

The attentional weight function $w_s$ is convex. This implies that when the treatment variable $\Delta$ increase, the attentional weight of the bad state $w_b$ will decrease faster (at a diminishing rate) than the attentional weight of the good state $w_g$. The effect is that the odds $\frac{\pi}{1-\pi}$ are greater than 1 when $\Delta < \Delta^0$, equal to 1 when $\Delta = \Delta^0$, and less than 1 when $\Delta > \Delta^0$. The implication of this is that for low values of $\Delta$ salient thinkers may bet nothing. This is because the bad state can be overweighted to a degree where a salient thinker is not willing to bet any amount. Now consider a risk-averse ($v'' < 0$) salient thinker who is willing to bet some amount. Increasing $\Delta$ will make the good state relatively more salient, which increases the slope of his strictly convex indifference curve. Equality in equation (4) is then restored by betting more and 'moving down' the indifference curve (see also figure 1).

In Appendix A we prove the following proposition for the general attentional weight function.

**Proposition 1.** *Sufficient conditions for an effect of salience are:*

*(1) nothing is bet for low levels of $\Delta$; and/or*

*(2) the chosen bet $\alpha^*$ and the treatment variable $\Delta$ are (a) jointly increasing up to a point where the good consequence is salient, beyond which (b) $\alpha^*$ decreases.*

These two conditions are, either together or apart, sufficient for salience to have an effect. They are however not necessary. If not-risk-averse ($v' \geq 0$) salient thinkers bet all

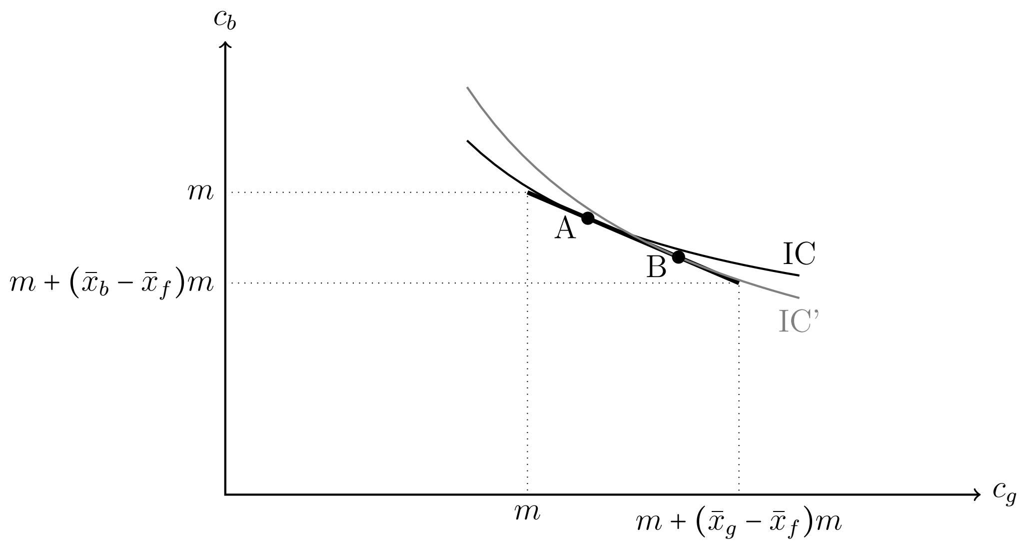


**Figure 1:** The bold black line segment defines fixed attainable wealth. IC and IC' represents the strictly convex indifference curves of a risk-averse salient thinker. Point A is an old optimum where IC is tangent to the bold black line. A relative increase in the salience of the good consequence will increase the slope of the indifference curve at A. IC' indicate this new indifference curve for which tangency can be restored by betting more until point B is reached.

their money, they are indistinguishable from not-risk-averse standard economic decision makers who are immune to salience distortions.

If attentional weights have the rank-dependent functional, then the following corollary follows from the proof of proposition 1.

**Corollary 1.** *When the attentional weight put on the state $s$ is $w_s = \delta^{k_s}$, where $\delta \in (0,1]$ and $k_s \in \{1,2\}$ with $k = 1$ indicating the more salient state, then point (2) in proposition 1 is replaced by:*

*(2') the amount bet is constant when the bad state is salient, and jump to a higher constant level when the good state is salient.*

Under rank-dependence the salience effect is therefore identified if $\alpha^*$ follows a step function with one positive jump at $\Delta = \Delta^0$.

Before presenting our experiment, note that Salience Theory for risky choice proposes two slightly different representations. The first describes how simple lotteries are compared, while the second lends itself to the analysis of asset allocation. The task subjects face in our experiment is a mixture of the two: they have to decide on the allocation of money between two simple lotteries.

Both characterizations assume that the value function is linear, and that salience weights are lottery/asset specific, which implies that the salience value is given by $V(\alpha) = v(e-m) + \sum_{s\in\{g,b\}} \pi_s^A v(x_s\alpha) + \pi_s^B v(x_f(m-\alpha))$ with $\pi_g^i + \pi_b^i = 1$ for $i \in \{A, B\}$ and $\pi_g^i = pw(\frac{x_g}{x_f}) \big/ [pw(\frac{x_g}{x_f}) + (1-p)w(\frac{x_f}{x_b})]$. Homogeneity of degree zero and the ordering property implies that $\pi_g^A = \pi_g^B \equiv \pi$. We can thus simplify to get our representation $V(\alpha) = \pi v(c_g) + (1-\pi)v(c_b)$. This representation is in line with Bordalo et al. (2012b). As an alternative, Bordalo et al. (2013a) suggest that salience is a function of the contrast between payoff $x_s$ and the average payoff $(x_s + x_f)/2$ in state $s$. In this case $\pi_s^A \neq \pi_s^B$, which implies that $V(\alpha) = \pi_g^A v(c_g) + (1-\pi_g^A)v(c_b)$. This representation does not change proposition 1.

By defining the value function on wealth consequences $c_s$ we follow Bordalo et al. (2013a). When choosing between two lotteries, however, Bordalo et al. (2012b) suppose that consequences (in their case payoffs) are evaluated relative to reference point zero. If this is the case, then subjects would choose the optimal amount to bet by computing, for each allocation the potential gains and losses relative to some reference wealth. We devote much of the next section to show that this representation does not confound our identification of the salience effect.

# 3 The Experiment

In this section we first describe the implementation of the experiment and show some descriptive statistics. Following this, we present the formal results relating to our hypothesis. We do so in two ways. First we present the results of a saturated Tobit regression analysis. Second, we back our results by quantifying the effect that salience has on decision weights using a structural model.

## 3.1 Data

The experiment was conducted on-line in November 2016 using the Internet Panel of the Center of Experimental Economics, University of Copenhagen. 1,300 panel members were invited by email to participate in the experiment. They were provided a link, a login and a password that could be used to participate. Upon logging on subjects were given instructions as well as control questions. After having answered all the control questions correctly they were allocated to one of the 24 possible sequences of the four different treatments (i.e. decision situations). More precisely, each subject was serially assigned to one of the 24 possible orders until all were exhausted and the process was repeated. This process implies that the order of the task presentation was fully balanced. In total 473 participants finished the complete experiment and generated 1,892 observations.

The treatment values were $\Delta = 0.0, 0.1, 0.4, 0.5$. Baseline payoffs were $\bar{x}_g = 1.1$, $\bar{x}_b = 0.1$, and $\bar{x}_f = 0.4$, and the probability of the good state obtaining was $p = 0.6$. These parameter values implied that subjects faced 4 treatment levels, labelled $T = 0, 1, 2, 3$, summarized in Table 1.

**Table 1:** Experimental Parameters

| | Treatment levels | | | |
|---|---|---|---|---|
| | $T = 0$ | $T = 1$ | $T = 2$ | $T = 3$ |
| *(in DKK)* | | | | |
| Endowment $e$ | 160 | 150 | 120 | 110 |
| Money to bet $m$ | 100 | 100 | 100 | 100 |
| *Lottery $L_A$:* | | | | |
| Payoff $x_g$ | 1.1 | 1.2 | 1.5 | 1.6 |
| Payoff $x_b$ | 0.1 | 0.2 | 0.5 | 0.6 |
| *Lottery $L_B$:* | | | | |
| Payoff $x_f$ | 0.4 | 0.5 | 0.8 | 0.9 |
| Probability $p$ | 0.6 | 0.6 | 0.6 | 0.6 |

The bad consequence was salient at treatment level 0 and treatment level 1 and the good consequence was salient at treatment level 2 and treatment level 3. The value of the treatment variable that implied equal salience was $\Delta^0 = 0.125$.[6]

In the instructions subjects were informed that payoffs would be expressed in actual Danish crowns (DKK), and that only one of the four treatments would randomly be chosen at the end of the experiment to be paid out. Subjects did not get any feedback regarding the outcomes during the experiment. They only got to know the outcome of one treatment that was randomly picked at the very end of the experiment. Payments were transferred to subjects' bank accounts within four weeks after the end of the ex-

[6] Had salience been defined relative to average state $s$ payoff, then the point of equal salience would have been located at $\Delta^0 = 0.3875$. Thus not changing the salience of states.

periment. Subjects earned on average DKK 118 (approx. USD 18). Screen-shots of the instructions and tasks are included in Appendix B.

**Table 2:** Summary of Experimental Data

| | Treatment levels | | | |
|---|---|---|---|---|
| | $T = 0$ | $T = 1$ | $T = 2$ | $T = 3$ |
| *Amount bet:* | | | | |
| Mean | 67.5 | 71.7 | 77.3 | 76.5 |
| 1st quantile | 50 | 50 | 60 | 60 |
| Median | 70 | 80 | 80 | 85 |
| 3rd quantile | 100 | 100 | 100 | 100 |
| *No. censored obs. at:* | | | | |
| Min (0) | 38 | 33 | 13 | 17 |
| Max (100) | 172 | 187 | 202 | 215 |

Across treatments the vast majority of subjects chose to bet some amount. Only 5 of the 473 subjects did not bet anything in the experiment. Table 2 presents a descriptive summary of the experimental data. The mean amount bet by subjects increases from treatment level 0 to treatment level 2, and slightly decreases from treatment level 2 to treatment level 3 (row 1). Judging by the quantiles it seems like the distribution shifts towards betting more when possible (rows 2-4). This is supported by the observation that, when going from treatment level 0 to treatment level 2, fewer subjects are censored by the minimum 0 and more by the maximum 100 (rows 5-6).

## 3.2 Treatment Effects

To test whether we in our experiment identify the salience effect we use following saturated Tobit regression model (censored at 0 and 100) based on the observed amount bet for each subject $i$ at each treatment level $j$:

$$\alpha^*_{i,j} = a + b_1 D_{i,1} + b_2 D_{i,2} + b_3 D_{i,3} + u_{i,j} \tag{5}$$

with $D_{i,j} = 1[T_i = j]$ being a dummy indicating treatment level $j$ and $u_{i,j} \sim N(0, \sigma)$. The zero-mean normal distributed error term $u_{i,j}$ represents unobserved factors other than the treatments that affect $\alpha^*_{i,j}$.

According to proposition 1 the salience effect is on the subject level identified if: (i) the amount bet at treatment level 1 is greater than zero, (ii) the amount bet at treatment level 2 is greater than that bet at treatment level 1, and (iii) the amount bet at treatment level 3 is at least greater than that bet at treatment level 1. Proposition 1 is, however, based on Salience Theory which is deterministic, while the observed data is stochastic in the sense that there exist extraneous variation at the level of each choice. In the saturated Tobit regression model this variation is 'averaged out' by the estimation procedure. Leaving the $j$th-level average treatment effect (ATE-$j$), given by the $j$th slope:

$$b_j = E[\alpha^*_{i,j} - \alpha^*_{i,0}],$$

with constant $a = E[\alpha^*_{i,0}]$ being the 'control level'.

A concern with this estimation could be that treatment effects may be confounded by the order of presentation in a way in which, the effect of early treatments could bias later treatments. However, our fully balanced design should on average cancel out this confound. A further concern may be that unobserved errors are correlated across treatments at the subject level. In other other words, the unobserved variance in a subject's choice may not be independent, but rather correlated due to some personal trait. If this concern is valid, then there exist a clustered error problem that must be taken care of. We address these concerns in the following estimates of the treatment effects.

Specification 1 of Table 3 estimates equation (5), controlling for within subjects correlations in the error terms using robust cluster errors. The estimated ATE-1, ATE-2, and ATE-3, are DKK 6.18, DKK 14.96, and DKK 15.08, respectively. All estimates are significantly different from zero. Specification 2 verifies that the fully balanced design worked. Controlling for the observed order of presentation has no effect on estimated ATEs, nor their significance. In specification 3 we estimate a model analogous to specification 1, but without the robust clustered errors. The estimates of this specification are the same as in our main specification 1, but ATE-1 is now insignificant. This is likely due to the the clustered standard error problem.

Table 3 shows that the salience effect is identified by the data. According to Proposition 1 it is sufficient for identification of salience that: (i) ATE-1 is greater than zero, (ii) ATE-2 is greater than ATE-1, and (iii) ATE-3 is greater than ATE-1. A t-test on specification 1 indicates that ATE-1 is not zero ($p = 0.014$). F-statistics indicate that

**Table 3:** Average Treatment Effects: Tobit Regression Estimates

| | (1) | (2) | (3) |
|---|---|---|---|
| Treatment level 1 | 6.18 | 6.18 | 6.18 |
| | (2.50) | (2.51) | (3.46) |
| Treatment level 2 | 14.96 | 14.96 | 14.96 |
| | (2.61) | (2.61) | (3.48) |
| Treatment level 3 | 15.08 | 15.08 | 15.08 |
| | (2.82) | (2.82) | (3.49) |
| Constant | 77.68 | 77.07 | 77.68 |
| | (2.63) | (4.17) | (2.45) |
| Order of treatment | - | 0.05 | - |
| | | (0.28) | |
| Robust cluster errors | Yes | Yes | No |
| Sample size | 1892 | 1892 | 1892 |
| Number of clusters | 473 | 473 | - |
| Log likelihood | -6157.1 | -6157.05 | -6157.1 |

*Notes:* All columns report estimates of a two-sided tobit regression model censored at 0 and 100. Of the 1,892 observations, 101 were left-censored and 766 right-censored. The coefficient estimates are all in terms of the underlying latent dependent variable. Specification 1 reports our main estimation model as specified in equation (5) with clustered standard errors. Specification 2 reports the the estimation model controlling for the order of presentation with clustered standard errors. Specification 3 reports plain Tobit regression estimates of the main estimation model. The Standard errors are reported in parentheses beneath coefficient estimates.

ATE-2 ($p = 0.0003$) and the ATE-3 ($p = 0.0016$) are both significantly different from the ATE-1.

Table 3 also shows that Corollary 1 cannot be supported by the data. If salience distortions were rank-dependent, then ATE-1 should be zero and ATE-2 should be equal to ATE-3. The first hypothesis is falsified by the t-test above, while the latter cannot be rejected ($p = 0.9603$). However, rank-dependent salience distortions imply that the two need to be jointly true.

## 3.3 A Structural Model

To go beyond the conclusion of the closed-form identification presented in the previous section, we estimate the effect that salience has on decision weights. We apply a structural model that defines how the observed bet amount relates to a salience parameter. Estimating such a parameter could aid in quantifying impacts on specific outcomes beyond that implied by our experimental design.

To this end, we represent the homothetic attentional weight by the following functional:

$$w(x,y) = \left(\frac{y}{x}\right)^{\gamma}.$$

This representation makes it cardinal, and restricting the attentional weight to the homeomorphic subset. The salience parameter $\gamma$ measure the curvature of the attentional weight function. Thus, the attentional weight parameter is increasingly concave for smaller values of $\gamma < 1$. This representation directly imposes the following linear relationship in the log odds metric:

$$\log\left(\frac{\pi}{1-\pi}\right) = \log\left(\frac{p}{1-p}\right) + \gamma\ \log\left(\frac{x_g/x_f}{x_f/x_b}\right).$$

The salience parameter $\gamma$ thus characterizes the elasticity of the perceived odds with respect to the relative ratio of payoffs. The decision weight $\pi$ is unfortunately not observable in the data.

We need to make some assumption about the stochastic structure underlying the observations. Salience Theory is of no help in this respect since it is a theory of deterministic choice. This, in turn, implies that the subject must be revealing his risk attitude with some error. We make the natural assumption that when searching for the optimal amount to bet, subjects adjust their marginal rate of substitution with some randomness. If these adjustments are made often enough, then the marginal rate of substitution that ensures optimal betting will be lognormal distributed, and we may write $\ln\left(\frac{v'(c_g^*)}{v'(c_b^*)}\right) + u_{i,j}$ with $u_{i,j}$ being the zero-mean normal distributed error term.[7]

[7] Let $v'(c_g^*)$ be denoted $v'(c_1)$, $v'(c_b^*)$ be denoted $v'(c_0)$, and index in sequence any marginal value between these two by the real line $]0,1[$. For any positive integer $m$, setting $h = 1/m$, $\frac{v'(c_1)}{v'(c_0)} = \frac{v'(c_1)}{v'(c_{(m-1)h})}\frac{v'(c_{(m-1)h})}{v'(c_{(m-2)h}))}\cdots\frac{v'(c_h)}{v'(c_0)}$. Consequently, taking logarithms, we get $\ln\left(\frac{v'(c_1)}{v'(c_0)}\right) = \sum_{k=1}^{m}\ln\left(\frac{v'(c_{kh})}{v'(c_{(k-1)h})}\right)$, where the factors $\frac{v'(c_{kh})}{v'(c_{(k-1)h}}$ are independent and identically distributed random variables. By the Central Limit Theorem, under mild additional conditions–for example, if $\ln\left(\frac{v'(c_1)}{v'(c_0)}\right)$ has finite variance, then

Taking into account the randomness of the marginal rate of substitution, the structural estimation will be based on the optimality condition stated in equation (4). This implies a general structural equation given by:

$$\ln\left(\frac{v'(c^*_{b,i,j})}{v'(c^*_{g,i,j})}\right) = \ln\left(\frac{x_g - x_f}{x_f - x_b}\right) + \ln\left(\frac{\pi}{1-\pi}\right) + u_{i,j}. \tag{6}$$

This equation jointly depends on the decision weight $\pi$ and the curvature of the value function $v$. It can therefore not be estimated without assuming some functional for $v$.

We consider both the constant absolute risk aversion (CARA) functional, $v(c) = 1 - e^{-\rho c}$, and the constant relative risk aversion (CRRA) functional, $v(c) = c^{1-\rho}/(1-\rho)$. In the first the characterization of risk aversion varies with wealth level, while in the second it is constant. This is per se not an issue in our experimental design where consequences are constant across treatment. However, by considering both we can evaluate the robustness of our salience parameter estimate. For observable outcome $y^*_{i,j} = \alpha^*_{i,j}$, we get the following CARA Tobit regression model (censored at 0 and 100):

$$y^*_{i,j} = a + b\ \log\left(\frac{x_{g,j}}{x_{f,j}} \Big/ \frac{x_{f,j}}{x_{b,j}}\right) + u_{i,j} \tag{7}$$

with $a = \log(3.5)/\rho$, $b = \gamma/\rho$, and $u_{i,j} \sim N(0, \sigma/\rho)$.[8] We can change regression equation (7) into a CRRA Tobit regression model (censored at 0 and 100) by instead considering observable outcome $y^*_{i,j} = \ln\left[\frac{c^*_{g,i,j}}{c^*_{b,i,j}}\right]$.

Based on our regression equation (7) and its two representations, Table 4 reports Tobit regression estimates with clustered standard errors. Specification 1 of Table 4 estimates the CARA specification. The estimated coefficient $\hat{b}$ is 28.11 (p=0.000) and the estimated constant $\hat{a}$ is 87.35 (p=0.000), both significantly different from zero. This implies that the jointly derived estimates are $\hat{\rho} = 0.014$ and $\hat{\gamma} = 0.39$, respectively. Increasing the relative payoff contrast by one percent will increase odds by 0.39 percent. Specification 2 of Table 4 estimates the CRRA specification. For this specification the estimated coefficient $\hat{b}$ is 0.13 (p=0.000) and the estimated constant $\hat{a}$ is 0.41 (p=0.000), both significantly different from zero. This implies that the jointly derived estimates

arbitrarily large $m$ implies that $\ln\left(\frac{v'(c_1)}{v'(c_0)}\right)$ is approximately normal distributed.

[8] First recognize that the CARA value function has the following representation: $\log(v') = -\rho c + d$ for some $d$. It is straightforward to solve for $\alpha^*$ in equation (4). Finally, apply the parametric values from the experiment and the estimation procedure from section 3.2.

**Table 4:** Structural Equations: Tobit Regression Estimates

| | (1) Amount bet | (2) Log relative wealth |
|---|---|---|
| Log ratio of relative payoffs | 28.11 | 0.13 |
| | (4.6) | (0.21) |
| Constant | 87.35 | 0.41 |
| | (2.17) | (0.01) |
| Robust cluster errors | Yes | Yes |
| Sample size | 1892 | 1892 |
| Number of clusters | 473 | 473 |
| Log likelihood | -6157.71 | -1341.17 |
| $\hat{\rho}$ | 0.014 | 3.06 |
| $\hat{\gamma}$ | 0.39 | 0.4 |

*Notes:* Both columns report estimates of a two-sided tobit regression model. The CARA specification is censored at 0 and 100, and the CRRA specification is censored at 0 and 0.46. Of the 1,892 observations, 101 were left-censored and 766 right-censored. The coefficient estimates are all in terms of the underlying latent dependent variable. Specification 1 reports our main CARA estimation model as specified in equation equation (7) with clustered errors. Specification 2 reports the CRRA estimation model with clustered standard errors. The derived estimates are calculated by $\hat{\rho} = \log(3.5)/\hat{a}$ and $\hat{\gamma} = \hat{\rho}\hat{b}$. Standard errors are reported in parentheses beneath coefficient estimates.

are $\hat{\rho} = 3.06$ and $\hat{\gamma} = 0.4$, respectively. This suggest that the estimation of the salience parameter $\gamma$ is robust across specifications.

# 4 Benartzi and Thaler's Model

As already hinted at in the introduction, Benartzi and Thaler (1995)'s adaptation of Prospect Theory to portfolio choice can also in principle explain that subjects allocate more to risky assets across the treatments, but to which extent can it numerically account for our experimental results?

Addressing the equity premium puzzle, Benartzi and Thaler propose that investors in asset markets judge their portfolios by the value of their holdings and not their overall wealth levels. Applying this intuition to our experimental setting means that the valuation of a prospect would be based on $x_s\alpha + x_f(m - \alpha)$, and not the wealth consequence $c_s$. A central tenet of Prospect Theory is that people derive value from gains and losses, measured relative to some reference point, rather than from absolute levels of wealth. This way of thinking implies that Benartzi and Thaler's valuation is based on $c_s - r$ with $r = e - m$ being the reference point. In our experiment this would imply that a subject calculates the difference between the endowment and the money available for betting, and that the subject implicitly or explicitly uses this amount as a reference point. Plugging in the baseline payoffs gives $x_s\alpha + x_f(m - \alpha) = (\bar{x}_f + \Delta)m + (\bar{x}_s - \bar{x}_f)\alpha$, so the value of the holding would be increasing in the treatment variable $\Delta$ and always positive.

Given this, a subject motivated by Prospect Theory à la Benartzi and Thaler (1995) in our experiment would evaluate consequences, and choose $\alpha$, the amount bet on the risky lottery $L_A$, to maximize

$$V(\alpha) = \pi(p)v(c_g - r) + (1 - \pi(p))v(c_b - r), \tag{8}$$

with $\pi$ now being Prospect Theory's decision weights.

The valuation $v$ function is according to Prospect Theory concave in the positive domain. The subject will bet an amount up to the point where $V'(\alpha) = 0$. By equation (8) we therefore have that

$$-\frac{x_f - x_b}{x_g - x_f} = -\frac{\pi(p)}{1 - \pi(p)} \frac{v'(c_g - r)}{v'(c_b - r)}. \tag{9}$$

As the treatment variable $\Delta$ increases, the value of the holding will also increase. By the concavity of the value function, $\frac{v'(c_g - r)}{v'(c_b - r)} < 1$ and converging towards 1 as the treatment variable $\Delta$ is increased. Increasing $\Delta$ will thus increase the slope of the strictly convex indifference curve. Equality in equation 9 is then restored by betting more thereby 'moving down' the indifference curve.

To assess the extent to which Benartzi and Thaler's model actually explains our observations, we modify our structural model (6):

$$\ln\left(\frac{v'(c^*_{b,i,j} - r)}{v'(c^*_{g,i,j} - r)}\right) = \ln\left(\frac{x_g - x_f}{x_f - x_b}\frac{\pi(p)}{1 - \pi(p)}\right) + u_{i,j}. \tag{10}$$

In our estimation we will consider the power functional $v(c_s - r) = (c_s - r)^\lambda$ with $0 < \lambda < 1$, which is customarily used in applications of Prospect Theory when $c_s - r$ is positive. For observable outcome $y^*_{i,j} = \ln\left[\frac{c^*_{g,i,j}-r_j}{c^*_{b,i,j}-r_j}\right]$, we apply the following Tobit model with variable censoring across treatments:

$$y^*_{i,j} = a + u_{i,j} \text{ with } a = \frac{1}{1-\lambda}\ln\left(3.5\frac{\pi(0.6)}{1-\pi(0.6)}\right). \tag{11}$$

The estimated constant is $\hat{a} = 1.399$ (p=0.000), which implies that:

$$\hat{\pi}(0.6) = \frac{e^{1.399(1-\lambda)}\frac{0.3}{0.7}}{1+e^{1.399(1-\lambda)}\frac{0.3}{0.7}}. \tag{12}$$

There are two ways to demonstrate that the estimated parameters are implausible. First, by applying a feasible estimated range for the $\lambda$-parameter in the value function $v$, $\hat{\lambda} \in [0.7, 0.9]$ (Tversky and Kahneman, 1992; Abdellaoui, 2000; Harrison and Rutström, 2009), we find by implication that $\hat{\pi}(0.6) \in [0.3, 0.4]$. But these studies all jointly estimate the $\lambda$-parameter and the weighting function

$$\pi(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{1/\gamma}},$$

and actually find that $\hat{\gamma} \in [0.6, 0.9]$. This would imply that $\hat{\pi}(0.6)$ should really lie in the range $[0.5, 0.6]$. The degree of underweighting we find necessary to account for our experimental data is therefore difficult to reconcile with the findings of these studies.

Second, assuming a more realistic underweighting $\hat{\pi}(0.6) = 0.5$ would by equation (12) imply that $\hat{\lambda} = 0.3$. That is, a representative subject would be indifferent between receiving approximately DKK 0.01 with certainty and a fifty-fifty win DKK 1 / lose DKK 0 lottery.

It is unlikely that most subjects underweight a probability of 0.6 by nearly two-thirds, or that most subjects are risk-averse to the degree that they would price such an Arrow-Debreu security at 0.01. Salience Theory offers in our view the more likely explanation.

# 5 Conclusion

This paper presents a direct experimental test of the prediction of Salience Theory, that the salience of the good consequence will make the risky lottery look more attractive. Our results strongly support this prediction. We manipulated the Weber-Fechner law by which subjects are thought to evaluate prospects. This manipulation was intended to change the salience of the consequences in our experimental design. As a consequence, salient thinkers will put disproportional weight on consequences when considering the value of prospects. In particular, we observe that subject bet more when good consequences are salient. The results provide support for the behavioral hypothesis of Salience Theory.

The results may also have relevance for asset pricing. After all, risky assets are lotteries evaluated in a context described by the alternative investments available in the market. A direct implication of our results is that the Weber-Fechner law causes investors to focus on downside risks more than on equal-sized upside risks, leading to an undervaluation of risky assets. This undervaluation will lead to lower prices, causing expected returns on the risky assets to increased. Our results thus also support a Salience Theory based explanation of the equity premium puzzle. Bordalo et al. (2013a) show that Salience Theory also can explain other puzzles in finance.

Of course, our experiment is highly stylized. For example, the subjects in the experiment only face known probabilities, whereas in real-life investors mainly deal with unknown probabilities. Another issue is that the financial stakes for the experimental subjects are low compared with those of most investors. Furthermore, it might be that trading wash out some the salience effect, if not all. These concerns are a cause for caution in extrapolating the results. However, the first two are not of major concern in so far our structural estimation of the salience parameter measures a 'deep psychological constant.' The latter concern suggest a line along which to pursue further experimental work.

# References


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

Allais, M., 1953. L'extension des théories de l'équilibre économique général et du rendement social au cas du risque. Econometrica, Journal of the Econometric Society, 269–290.

Barberis, N., Huang, M., 2008. Stocks as lotteries: The implications of probability weighting for security prices. American Economic Review 98 (5), 2066–2100.

Bell, D. E., 1982. Regret in decision making under uncertainty. Operations research 30 (5), 961–981.

Bell, D. E., 1985. Disappointment in decision making under uncertainty. Operations research 33 (1), 1–27.

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

Bordalo, P., Gennaioli, N., Shleifer, A., 2012a. Salience in experimental tests of the endowment effect. The American Economic Review 102 (3), 47–52.

Bordalo, P., Gennaioli, N., Shleifer, A., 2012b. Salience theory of choice under risk. Quarterly journal of economics 127 (3), 1243–1285.

Bordalo, P., Gennaioli, N., Shleifer, A., 2013a. Salience and asset prices. The American Economic Review 103 (3), 623.

Bordalo, P., Gennaioli, N., Shleifer, A., 2013b. Salience and consumer choice. Journal of Political Economy 121 (5), 803–843.

Bordalo, P., Gennaioli, N., Shleifer, A., 2015. Salience theory of judicial decisions. The Journal of Legal Studies 44 (S1), S7–S33.

Bordalo, P., Gennaioli, N., Shleifer, A., 2016. Competition for attention. The Review of Economic Studies 83 (2), 481–513.

Bordalo, P., Gennaioli, N., Shleifer, A., 2017. Memory, attention, and choice. Tech. rep., National Bureau of Economic Research.

Chew, S. H., 1983. A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the allais paradox. Econometrica, 1065–1092.

Dertwinkel-Kalt, M., Köhler, K., Lange, M. R., Wenzel, T., 2017. Demand shifts due to salience effects: Experimental evidence. Journal of the European Economic Association 15 (3), 626–653.

Dzhafarov, E. N., 2001. Fechnerian psychophysics. In: Smelser, N. J., Baltes, P. B. (Eds.), International encyclopedia of the social and behavioral sciences, 8th Edition. Pergamon Press, New York, pp. 5437–5440.

Dzhafarov, E. N., Colonius, H., 2011. The fechnerian idea. American Journal of Psychology 124 (2), 127–140.

Fishburn, P. C., 1983. Transitive measurable utility. Journal of Economic Theory 31 (2), 293–317.

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

Gul, F., 1991. A theory of disappointment aversion. Econometrica, 667–686.

Harrison, G. W., Rutström, E. E., 2009. Expected utility theory and prospect theory: One wedding and a decent funeral. Experimental Economics 12 (2), 133–158.

Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. Econometrica: Journal of the econometric society, 263–291.

Kontek, K., 2016. A critical note on salience theory of choice under risk. Economics Letters 149, 168–171.

Kőszegi, B., Rabin, M., 2007. Reference-dependent risk attitudes. American Economic Review 97 (4), 1047–1073.

Kőszegi, B., Szeidl, A., 2012. A model of focusing in economic choice. The Quarterly journal of economics 128 (1), 53–104.

Leland, J. W., 1994. Generalized similarity judgments: An alternative explanation for choice anomalies. Journal of Risk and Uncertainty 9 (2), 151–172.

Loomes, G., Sugden, R., 1982. Regret theory: An alternative theory of rational choice under uncertainty. The economic journal 92 (368), 805–824.

Loomes, G., Sugden, R., 1986. Disappointment and dynamic consistency in choice under uncertainty. The Review of Economic Studies 53 (2), 271–282.

Machina, M. J., 1982. "expected utility" analysis without the independence axiom. Econometrica, 277–323.

Rubinstein, A., 1988. Similarity and decision-making under risk (is there a utility theory resolution to the allais paradox?). Journal of economic theory 46 (1), 145–153.

Savage, L. J., 1954. The foundations of statistics. John Wiley & Sons.

Schwartzstein, J., 2014. Selective attention and learning. Journal of the European Economic Association 12 (6), 1423–1452.

Tversky, A., Kahneman, D., 1992. Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and uncertainty 5 (4), 297–323.

Von Neumann, J., Morgenstern, O., 1944. Theory of games and economic behavior. Princeton university press.

Woodford, M., 2012. Prospect theory as efficient perceptual distortion. The American Economic Review 102 (3), 41–46.

# Appendix A: Proofs

## Proof of Proposition 1

(1). A salient thinker will not bet any amount if $V'(0) = v'(e - m + x_f m)[\pi x_g + (1 - \pi)x_b - x_f] \leq 0$, which is true if, and only if, the salience distorted risk premium $\pi x_g + (1-\pi)x_b - x_f]$ is positive. Increasing $\alpha$ from zero will not increase $V(\alpha)$ and $\alpha^* = 0$ will maximize his salience value. Absent salience distortions, the risk premium is by design always positive. However, with salience distortion, he may perceive the risk premium to be zero or negative. This happens when $w\left(\frac{\bar{x}_g+\Delta}{\bar{x}_f+\Delta}\right)\Big/w\left(\frac{\bar{x}_f+\Delta}{\bar{x}_b+\Delta}\right) \leq \frac{1-p}{p}\frac{\bar{x}_f-\bar{x}_b}{\bar{x}_g-\bar{x}_b}$ which obtain if and only if $\Delta \leq \tilde{\Delta} < \Delta^0$, where $\tilde{\Delta}$ is the level of the treatment variable that ensures equality in the expression. The risk premium can only be perceived as negative for low levels of $\Delta$ where bad consequence is salient.

(2). If a salient thinker perceives the risk premium as positive, he will always be bet some amount since $[\pi x_g + (1-\pi)x_b - x_f] > 0$ implies $V'(0) > 0$. How much will depend on shape of the value function. A "not-risk-averse" salient thinker with convex ($v''(\cdot) \geq 0$) value functions will always increase $V(\alpha)$ by betting more. Betting as much as possible will thus maximize his salience value. Without salience distortion it would still be optimal to bet as much as possible, so salience theory predicts zero effect for not-risk-averse betting.

It is left to analyze how the optimal bet depends on $\Delta$ for risk-averse salient thinkers with strictly concave ($v''(\cdot) < 0$) value functions. When the value function is strictly concave the indifference curve will be downward sloping and strictly convex. If the difference in attentional weights $D \equiv w\left(\frac{\bar{x}_g+\Delta}{\bar{x}_f+\Delta}\right) - w\left(\frac{\bar{x}_f+\Delta}{\bar{x}_b+\Delta}\right)$ increase as $\Delta$ increase, then (by Equation 4) the slope of the indifference curve will also increase and via MRS tangency is ensured by betting more. Conversely, the optimal amount to bet will decrease if $D$ decrease as $\Delta$ increase. Differentiating $D$ with respect to $\Delta$ gives

$$\left(\frac{\bar{x}_f+\Delta}{\bar{x}_b+\Delta}\right)^2 w'\left(\frac{\bar{x}_f+\Delta}{\bar{x}_b+\Delta}\right) - \frac{\bar{x}_g-\bar{x}_f}{\bar{x}_f-v_b}w'\left(\frac{\bar{x}_g+\Delta}{\bar{x}_f+\Delta}\right).$$

The first term of this expression is monotonically decreasing in $\Delta$, and so is the second term. When the bad consequence is salient, their difference is increasing in $\Delta$ since $w'\left(\frac{\bar{x}_f+\Delta}{\bar{x}_b+\Delta}\right) > w'\left(\frac{\bar{x}_g+\Delta}{\bar{x}_f+\Delta}\right)$ due to convexity of $w$, and $\left(\frac{\bar{x}_f+\Delta}{\bar{x}_b+\Delta}\right)^2 > \frac{\bar{x}_g-\bar{x}_f}{\bar{x}_f-v_b}$ when $\Delta < \Delta^0$. At $\Delta = \Delta^0$, where there is no salience distortion, the expression is still increasing since $w'\left(\frac{\bar{x}_f+\Delta^0}{\bar{x}_b+\Delta^0}\right) = w'\left(\frac{\bar{x}_g+\Delta^0}{\bar{x}_f+\Delta^0}\right)$ and $\left(\frac{\bar{x}_f+\Delta^0}{\bar{x}_b+\Delta^0}\right)^2 > \frac{\bar{x}_g-\bar{x}_f}{\bar{x}_f-v_b}$. The expression will continue to increase in $\Delta$ until an unique maximum is reached, beyond which it will decrease towards an asymptote defined by the non-salient amount bet. This maximum exist since $\Delta \to \infty$ implies that the expression at some point will be negative. Monotonicity ensures a global maximum.

# Appendix B: Screenshot

**CEE Virtual Laboratory**

**Round 1**

You start this round with 160 DKK. 100 DKK of these 160 DKK can be bet on Lottery A and B. Below you are shown the table with the information regarding 'Lottery A' and 'Lottery B' in this round:

| | Lottery A | Lottery B |
|---|---|---|
| 60% | 1.1 | 0.4 |
| 40% | 0.1 | 0.4 |

In this round you have a 60% chance to earn 1.1 times the amount you bet on Lottery A and a 40% chance to earn 0.1 times the amount you bet on Lottery A. The amount not bet on Lottery A will automatically be bet on Lottery B. You earn 0.4 times the amount bet on Lottery B in this round.

You will also receive the difference between the starting amount and the amount that you can bet, i.e 60 DKK, in this round.

Please decide how much of the 100 DKK you want to bet on Lottery A in this round.

**Your decision**

The amount (in DKK) that I would like to bet on Lottery A in this round is: [ ]

Press "Next" to continue.

**Next**

**Figure 2:** This is a screenshot of the decision situation at treatment level '0'. The other decision situations were identical beside the starting endowment and the payoffs of the lotteries. The starting endowment and the payoffs at the other treatment levels can be found in Table 1.