Testing for salience effects in choices under risk*

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

March 12, 2021

Abstract

We construct and run an experiment to test the most basic choice effect predicted by Salience Theory. Subjects allocate wealth between a risky and a safe investment. While we vary an apparent payoff ratio to influence salience, treatments have economically equivalent consequences. Most other theories of behavior then predict zero effect. Our experimental findings are strongly consistent with the behavioral implication of a continuous version of Salience Theory. We provide a novel structural estimate on the strength of salience. In our setting, increasing the relative payoff contrast by one percent is equivalent to an increased odds ratio by about 0.4 percent.

Keywords: Salience, attention, risky behavior

JEL-Classifications: D81, D91, G11, G41

*We are grateful for comments from Pedro Bordalo, Markus Dertwinkel-Kalt, Pietro Ortoleva and Adam Szeidl, as well as seminar and workshop participants at Aarhus University, Bocconi, University of Cologne, University of Copenhagen, Copenhagen Business School, ECARES (Bruxelles), European Economic Association Congress 2020, University of St. Gallen, Université Laval (Québec City), Lund University, NYU (Abu Dhabi), University of Southern Denmark (Odense), LMU Munich as well as Nice. 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: peter.norman.sorensen@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, attending to all available information. However, our cognitive resources are actually bounded, so we must often limit our attention when evaluating the many pieces of information we are confronted with. Once attention is limited, it can be drawn to salient features of our environment.[1] Starting from basic insights on human perception, Bordalo, Gennaioli, and Shleifer have recently developed Salience Theory to explore economic consequences of the way salience captures attention (Bordalo et al., 2012a,b, 2013a,b, 2015, 2016).

We design and run an experiment to test the essential choice effect predicted by Salience Theory. Salience Theory rests on two psychophysical concepts of perception, which in combination have become known as the Weber-Fechner law (Dzhafarov, 2001; Dzhafarov and Colonius, 2011). The first part of 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 as similar to a difference between 22 and 20. The second part is *Fechnerian sensitivity*. This postulates that there is diminishing sensitivity to a given difference in stimuli magnitude. Human perception is more sensitive to a difference between 11 and 10 than to a difference between 16 and 15.

In our experiment, individuals are asked to make some simple choices of allocating wealth to a risky investment. We design the experiment such that choices have economically equivalent consequences across treatments — the majority of theories of choice under risk thus predict zero treatment effect. Yet, we vary an apparent payoff ratio across treatments in order to potentially influence the salience of the risky choice. Salience Theory predicts a choice effect, which is thus isolated from most other theories for choices under risk. The experiment allows for the detection of this diminishing sensitivity effect, the core prediction of Salience Theory.

Let us illustrate the main idea of our experiment by explaining two of our treatments, A and B.[2] In treatments A and B subjects are respectively endowed with DKK 160

[1] See e.g. Kőszegi and Szeidl (2013), Bordalo et al. (2012b), Woodford (2012), Schwartzstein (2014), Dertwinkel-Kalt et al. (2017), Dertwinkel-Kalt and Köster (2017), Bordalo et al. (2017).

[2] Denoted A and B here in the introduction, these treatments are respectively denoted 0 and 3 in our later analysis.

Table 1: Treatment Example (all values in DKK)

	A	B
Endowment:	160	110
Money to bet:	100	100
Risky lottery:		
Probability: 0.6	1.1	1.6
Probability: 0.4	0.1	0.6
Risk-free lottery:	0.4	0.9

and DKK 110.[3] In both treatments, subjects can bet between DKK 0 and 100 of the endowment on a risky lottery. The residual part of the endowment that cannot be bet keeps its value (e.g., DKK 60 in treatment A).

[Table 1 here.]

Table 1 details the payoff numbers in the two treatments. In treatment A, the risky lottery gives a 60% chance to receive DKK 1.1 per DKK bet, and complementary chance 40% to receive DKK 0.1 per DKK bet. The remainder of the DKK 100, not bet on the risky lottery, is allocated to another, risk-free lottery that pays DKK 0.4 per DKK. The second column provides the corresponding information for treatment B.

This set-up is deliberately created such that the two treatments provide identical state-dependent wealth consequences to any amount bet on the risky lottery. If subjects bet DKK y on the risky lottery, then in both treatments the final payoff is DKK $100+0.7y$ with probability 60% and DKK $100-0.3y$ with probability 40%.[4] In other words, the prospects are exactly the same given any bet, and the choice set of bets from 0 to 100 DKK is identical in both treatments.

Salience Theory predicts the effect that subjects will bet more on the risky lottery in treatment B.[5] In application to risky decision making (Bordalo et al., 2012b, 2013a), consequences are described by payoff numbers. The relative payoffs to the risky bet differ in our treatments. By the Weber-Fechner law, the bad consequence is most salient in treatment A since $\frac{1.1}{0.4} < \frac{0.4}{0.1}$, while the good consequence is most salient in treatment B

[3] Danish Kroner (DKK) 160 was around USD 24 when the experiment was run.

[4] In Treatment A, the attained wealth will be DKK $60+1.1y+0.4(100-y)$ in the good consequence and DKK $60+0.1y+0.4(100-y)$ if the bad consequence obtains. In Treatment B, it will be DKK $10+1.6y+0.9(100-y)$ and DKK $10+0.6y+0.9(100-y)$, respectively.

[5] Section 2.4 discusses this effect in greater detail.

since $\frac{1.6}{0.9} > \frac{0.9}{0.6}$. More salient consequences get assigned larger decision weights in Salience Theory. In treatment B, the good consequence is more salient and weighted, causing risky betting to seem more attractive.

Clearly, in order to design the experiment in such a way that the two treatments provide identical state-dependent wealth consequences, we not only change relative payoffs across treatments in our set-up, but also the initial endowment. As explained in detail below, providing identical state-dependent wealth consequences across treatments allows us to control for many alternative theories, but changing the initial endowment might also invoke the house money effect Thaler and Johnson (1990). The house money effect says that people are more willing to take risks after a recent gain. That is, if participants in our experiment perceive the original endowment as 'house money', they should invest less into the risky lottery going from treatment level A to B. The opposite of our salience hypothesis. Hence, the salience hypothesis that we develop on the basis of (Bordalo et al., 2012b, 2013a) stands in sharp contrast to the house money effect in the context of our experimental set-up. As a consequence, if the house money effect is present in our data, then our estimates of the real underlying salience effect represent a lower bound.

To run this experiment, we invited subjects to a sequence of four treatments structurally similar to the illustration above. The four treatments varied 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 were equally salient, in theory. Participants were serially assigned to one of 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 already 77% of their DKK 100 on the risky lottery at baseline level 0.[6] Confirming the main prediction of Salience Theory, the average bet further increases by 18% towards treatment level 3.

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

[6]The numbers from Table 1 imply that the risky lottery offers an expected risk premium of 30%.

implied by our experimental design. We find that increasing the payoff contrast[7] by one percent, subjects respond as if the ratio of lottery probabilities (i.e., $\frac{0.6}{0.4}$) were increased by 0.39%.

To interpret these findings, let us turn to discuss the broader literature on risky choice behavior. As already mentioned, maintaining identical economic consequences across treatments, our experimental set-up controls for many existing theories of choice under risk. 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). The Focusing Theory of Kőszegi and Szeidl (2013) likewise distinguishes itself from Salience Theory when it requires that decision weights depend on the economic consequences (consumption or income) in each state, and is hence also mute here.[8]

It is also impossible for common consequence and common ratio effects (Allais, 1953) to explain outcome differences across our treatments. The common consequence effect would need at least a third consequence, which we do not have. The common ratio effect would need to change probabilities by proportion, but we kept probabilities fixed. The two-consequence feature of our experimental design also controls for more general theories that relax (Hong, 1983; Fishburn, 1983), or even discard (Machina, 1982), the 'independence axiom' of Expected Utility Theory.

Another obvious candidate theory is Prospect Theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992). Salience Theory shares with Prospect Theory a form of 'narrow framing,' which occurs when people evaluate a risk separably from other concurrent risks. Separating the two theories, Salience Theory predicts that probabilities are biased by payoffs (via the Weber-Fechner law), whereas in Prospect Theory value is derived from 'gains' and 'losses' measured relative to some reference wealth level. To apply Prospect Theory, an assessment must be made of the relevant reference point. The theory provides no unique guidance on how to find this reference point in general, but our experimental design controls for two significant suggestions.

[7]The payoff contrast in favor of the good consequence is the ratio $\frac{1.1}{0.4}/\frac{0.4}{0.1}$ in treatment A, for instance.

[8]Kőszegi and Szeidl comment that Focusing and Salience theories lead to similar intuitions, and that Focusing shares with Regret theory the property that "states in which the difference between lotteries is greater carry a greater weight." Again, Salience theory is distinct because the "difference between lotteries" is not defined by the ultimate payoff, but by some numbers that attract attention.

First, Barberis and Huang (2008) propose that investors perceive gains and losses relative to the wealth level they obtain from a risk-free investment.[9] In our setting this definition corresponds to the wealth obtained by subjects who bet nothing on the risky lottery. 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., DKK 60 in A and 10 in B, they receive the risk-free earnings DKK 40 and 90 respectively. The reference wealth level is thus DKK 100 in both treatments, and therefore cannot explain a treatment effect on betting behavior.[10]

Second, Kőszegi and Rabin (2006, 2007) propose that people rationally forecast potential wealth consequences, and derive utilities from the difference between actual consequences and these stochastic reference outcomes. Rationality implies that subjects' expectations will match the distribution of consequences they face prior to deciding on the amount to bet. In our example, a rational decision maker will perceive all treatments as identical, and hence hold identical reference beliefs. Since also consequences are identical across treatments, their theory is again unable to explain any change in behavior.

Nevertheless, 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.

It is important for us to point out that in designing our experiment and developing our hypotheses we exclusively focused on existing economic theories for choices under risk. This allowed us to clearly carve out the behavioral effects postulated by Salience Theory. Our experimental design cannot possibly rule out that new variants of these existing theories could also explain the behavioral effects we observe. Future theoretical

[9] Barberis and Huang (2008, p.2071): "This specification is more tractable and potentially more plausible: the investor thinks of the change in his wealth as a gain only if it exceeds the change he would have experienced by investing at the risk-free rate."

[10] Note that we take rationality to imply that subjects integrate their non-bettable wealth into the prospect.

and experimental analyses may further explore this possibility.

Experimental studies directly investigating Salience Theory and risk taking are few and en masse focus on settings where the predictions of Salience Theory contrast with those of Prospect Theory.[11] Bordalo et al. (2012b) provide suggestive evidence in favor of Salience Theory controlling for Prospect Theory. In their theory decision weights are determined by the salience of payoffs in a state, not the actual probabilities of the states, while decision weights in Prospect Theory are influenced by the probabilities. In unincentivized survey experiments conducted using Amazon Mechanical Turk, they find that Salience Theory successfully can explain survey data on the Allais common consequence and common ratio effect problems, as well as on prospects that differ by a mean-preserving spread.

Three other related papers that provide further experimental evidence for salience effects in choices under risk are Booth and Nolen (2012), Frydman and Mormann (2017) and Dertwinkel-Kalt and Köster (2019).[12] Also in these experimental analyses, the major focus of the employed designs lies in the differentiation of Salience Theory-based and Prospect Theory-based explanations for behavior. Booth and Nolen (2012) experimentally analyze gender differences in risk taking behavior in a simple hypothetical choice experiment originally presented by Bordalo et al. (2012b). Controlling for Prospect Theory they find that salience affects young men and women differently. Frydman and Mormann (2017) use a lottery choice problem commonly used in experiments on the Allais paradox and in this way successfully replicate the unincentivized Allais common consequence and common ratio effect problem implemented by Bordalo et al. (2012b). Controlling for Prospect Theory-based explanations, Dertwinkel-Kalt and Köster (2019) experimentally show that Salience Theory can explain skewness preferences.

[11]Two studies that analyse salience theory in situations not involving any risk are Dertwinkel-Kalt et al. (2017) and Dertwinkel-Kalt and Köster (2017). These analyses provide laboratory evidence consistent with Salience Theory.

[12]Königsheim et al. (2019) use a multiple price list approach to parameterize and estimate salience effects. Instead of controlling for other theories by design, they use a finite mixture approach to estimate the share of people that behave consistently with a rank-dependent version of Salience Theory and estimate a salience parameter which is broadly in line with the estimate of Bordalo et al. (2012b). As noted by the authors ((Königsheim et al., 2019, p. 488)), alternative decision theories like expected utility theory or rank-dependent utility theory 'also do a good job in explaining subjects' behavior' in their experiment. Furthermore, Lian et al. (2019) also experiment with subjects allocating wealth between a risky and a risk-free asset in a realistic setting. Their treatment reduces all return realizations by a fixed difference rate. Consistent with Salience Theory, subjects allocate more to the risky asset when rates are thus reduced. They note that expected-utility maximizers with decreasing absolute risk aversion should exhibit the opposite reaction to their treatment, but they do not control for other context-dependent choice models.

Although already discussed in the original analysis by Bordalo et al. (2012b, pp. 1258–1259), alternative explanations for behavior have not received much attention in the hitherto existing experimental analyses of salience affects. Bordalo et al. (2012b), for example, note that similarity-based preferences (e.g. (Rubinstein, 1988; Leland, 1994)) and regret theory (e.g. (Bell, 1982; Loomes and Sugden, 1982)) might also be able to explain several patterns of choices under risk consistent with the predictions of Salience Theory.[13]

In contrast to the existing experimental designs, the focus of our analysis is on creating a simple experimental design which not only allows to control for Prospect Theory but also for an encompassing range of alternative choice theories (see the aforementioned discussion on how our design controls for Expected Utility Theory, different forms of Prospect Theory, Regret Theory, Disappointment Theory, and Similarity-Based Theories). In addition, our analysis aims at creating an experimental choice paradigm which allows to develop unique theoretical predictions for Salience Theory without making restrictive assumption regarding e.g. the salience function or people's risk preferences.[14] Given this, our experimental findings contribute the first evidence consistent with predictions of Salience Theory while ruling out that other context-dependent choice models can explain the behavioral findings.

Our experiment isolates the core mechanism in Salience Theory using a very simple experimental setting, better controlling for alternative explanations. This not only allows us to derive unique theoretical predictions for salience with only minimal restrictions on the salience function, but also allows us to better quantify the impact of the payoff contrast on the perceived odds. The simplicity of our setting speaks more to the internal validity of Salience Theory than its external validity. When economic agents make actual choices under risk, exogenous variation would rarely manipulate salience without affecting other things. We take advantage of the laboratory setting where this variation is under our control.

Our analysis is organized as follows. The next section explains and motivates the design of our 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.

[13] To rule out such effects, Frydman and Mormann (2017) report evidence on the binary choice between a lottery or a safe outcome. They argue that their observed within-subject behavior makes it less likely that regret theory explains their findings, although it remains logically possible. A similar conclusion is reached by Dertwinkel-Kalt and Köster (2019).

[14] Frydman and Mormann (2017) and Dertwinkel-Kalt and Köster (2019) develop their theoretical predictions under the assumption that subjects are risk-neutral.

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. We derive salience effects under both the rank-dependent and continuous versions of Salience Theory.

2.1 Task

In each of four decision situations in our experiment, subjects are given an endowment e and a budget $m \leq e$. They are asked which amount $y \leq m$ they want to bet on a risky lottery L_A. If y is chosen, subjects face chance p of realizing a 'good' outcome with lottery earnings $x_g y$ and chance $1-p$ of realizing a 'bad' outcome with lottery earnings $x_b y$. The amount not bet on L_A is automatically allocated to a risk-free lottery L_B that pays $x_f(m-y)$ 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.

The state-dependent wealth consequence of choosing y is

$$c_s = e - m + x_s y + x_f(m-y) > 0, \tag{1}$$

with $s \in \{g, b\}$. Subjects that do not bet anything (setting $y = m$), are certain to earn the amount $x_f m$ from their choice. 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).

In our experiment, we vary the relative salience of the different outcomes while keeping prospects fixed. Precisely, we set payoffs $(x_g, x_b, x_f) = (\bar{x}_g + \Delta, \bar{x}_b + \Delta, \bar{x}_f + \Delta)$, with $\bar{x} > 0$ being a baseline, and $\Delta \geq 0$ a treatment variable.[15] We also adjust the endowment as $e = \bar{e} - \Delta m$. We keep budget m and probability p constant. In order to maintain $e \geq m$, we restrict the treatment variable to $\Delta \leq \Delta^U$, where, by definition, $\Delta^U = \frac{\bar{e}-m}{m}$. It is easy to verify from (1) that, for any $\Delta \in [0, \Delta^U]$, choice y provides identical prospect $(c_g, p; c_b, 1-p)$.[16]

[15]Actual numbers are in Table 2 below.

[16]Namely, $c_s = e-m+x_s y+x_f(m-y) = \bar{e}-\Delta m-m+\bar{x}_s y+\Delta y+\bar{x}_f(m-y)+\Delta(m-y) = \bar{e}-m+\bar{x}_s y+\bar{x}_f(m-y)$.

2.2 Salience Theory

First, Salience Theory for risky decision making as characterized by Bordalo et al. (2012b, 2013a) defines the salience of a state as a function of the possible payoffs that can be obtained in that state. The two articles propose slightly different implementations of the idea. The first article considers discrete choice in a finite set of lotteries. The second considers the continuous choice of allocating wealth to a finite set of assets. The choice problem in our experiment is a mixture of the two: subjects have to decide on the allocation of money between two simple lotteries. We adapt from Bordalo et al. (2013a) that salience is a function of the list of payoffs from the finite set of assets — here, our two lotteries. We will not literally follow Bordalo et al. (2012b) when they also assume that salience can be attached to a specific asset in any state. Here, instead, we stay closer to Bordalo et al. (2013a).[17]

Salience is then a function of the return pair $x, z > 0$ given in any particular state via the two available lotteries.[18] The salience function satisfies two main properties: ordering and diminishing sensitivity. Ordering says that, if an interval $[x, z]$ is contained in a larger interval $[x', z']$, then the pair (x, z) is less salient than the pair (x', z'). Diminishing sensitivity says that $(x + \epsilon, z + \epsilon)$ is less salient than (x, z) for all $x, z, \epsilon > 0$.

We follow Bordalo et al. (2013a) by also assuming that the salience function is symmetric and homogeneous of degree zero: (tx, tz) and (x, z) and (z, x) are equally salient for all $x, z, t > 0$.[19] Salience is thus defined as a function of the positive ratio $r = z/x$ for any pair $x, z > 0$. We let σ denote this salience as a function of ratio $r > 0$. The salience of (x, z) is then $\sigma(z/x)$. By symmetry, $\sigma(r) = \sigma(1/r)$. Ordering means that $\sigma(r)$ increases as r is further from 1. Diminishing sensitivity follows from these assumptions.[20]

Salience can now, in principle, be described by any increasing function $\sigma(r)$ defined for $r \geq 1$. Then, for $r < 1$, let $\sigma(r) = \sigma(1/r)$. The salience function reflects Weber-Fechner's stylized facts of human perception. Our perceptive apparatus is attuned to

[17] Bordalo et al. (2013a) suggest that salience is a function of the contrast between one asset's payoff x_s and the assets' average payoff $(x_s + x_f)/2$ in state s. With only two lotteries, salience is then a function of the pair (x_s, x_f), as we will assume.

[18] Negative payoffs can also be accommodated, but this is irrelevant for our experimental design.

[19] In recent work, Lanzani (2020) axiomatizes Salience Theory for binary choice settings. This leads to the conclusion that ordering is the defining property that distinguishes Salience Theory from expected utility or Prospect Theory, while diminishing sensitivity combines with a property of reflexivity to provide risk aversion in gains and risk love in losses.

[20] To see this, take any $x, z, \epsilon > 0$. By symmetry, it is without loss of generality to assume $z > x$. Then $\frac{z}{x} > \frac{z+\epsilon}{x+\epsilon} > 1$, so the salience of (x, z) exceeds the salience of $(x + \epsilon, z + \epsilon)$.

increases in contrast, here measured by increases in the distance of the payoff ratio r from one. As a famous example, Fechner (1860) proposes $\sigma(r) = K|\ln(r)|$ where $K > 0$ is a constant. For another prominent example, Bordalo et al. (2013a) suggest salience function $\frac{|x-z|}{x+z}$, which here becomes $\sigma(r) = \frac{|r-1|}{1+r}$. Below, we will also consider salience functional forms $(r-1)^\gamma$ and r^γ for $r \geq 1$, where $\gamma > 0$ is a parameter.

Second, salience σ does not directly affect decision weights, but works through attentional weights w_s which are functions of states' salience. When deciding, the salient thinker is presumed to weight the good state by:

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

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

In the rank-dependent version of Salience Theory, the attentional weight put on any state s is $w_s = \delta^{k_s}$. Here, $k_s \in \{1,2\}$ is the salience ranking of the two states, with $k_s = 1$ indicating the more salient state s — this is the state where contrast x_s/x_f is furthest from 1. If the two states are equally salient, they obtain the same ranking. The parameter $\delta \in (0,1]$ measures the extent to which salience distorts valuations. The salient thinker thus distorts the more and less salient state by δ and δ^2, respectively.[21] When $\delta = 1$, there is no salience distortion, but as δ is closer to zero, the salient thinker focuses ever more attention on the more salient state.

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

Below, we focus mainly on the continuous version of Salience Theory. Here, the attentional weight on state s is an increasing function of its salience σ_s. Since salience is itself an increasing function of ratio $r > 1$, it is without loss of generality to let $w_g = \sigma(\frac{x_g}{x_f})$ and $w_b = \sigma(\frac{x_f}{x_b})$ in the analysis of our experiment. The resulting decision weight attached to the good state is given by equation (2).

Third and finally, the salient thinker uses increasing value function v to evaluate

[21] What matters is the relative weight $\delta^2/\delta = \delta < 1$ attached to the less salient state.

realized payoffs, and chooses amount y to bet on the risky lottery in order to maximize

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

The attitude towards risk as measured by the optimal amount y^* 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 π. We will refer to the latter as the 'salience effect'. This 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, framing in Prospect Theory. In Prospect Theory, narrow framing implies that consequences are perceived as payoff gains and losses relative to a reference point.

2.3 Comparative Statics in Decision Weights

Given the treatment, a salient thinker maximizes $V(y)$ over $y \in [0, m]$. This maximization problem is straightforward, after inserting equation (1) into (3). If the solution is interior to the constraint set $[0, m]$, the optimal amount y^* must satisfy the usual first-order condition,

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

The left-hand side is the rate at which attainable wealth can be exchanged between the two outcome states (slope of the budget line). The right-hand side defines the rate at which a salient thinker is willing to take on risk (slope of the indifference curve). It may occur that the optimum is at an end of the constraint set $[0, m]$, corresponding to an end point of the budget line. In that case, (4) holds as an inequality.

By our construction, $x_f - x_b$ and $x_g - x_f$ are not affected by treatment variable Δ. Thus, the left-hand side of (4) is unaffected by treatment — this is also obvious from the fact that this is the slope of the budget line, and the budget set was deliberately held constant. On the right-hand side, the only exogenous variable to change with treatment is the ratio $\frac{w_g}{w_b}$. Observed changes in endogenous variable y must stem from this change. In line with intuition, the greater is the relative salience of the good state, the greater is predicted choice y.

[Figure 1 here.]

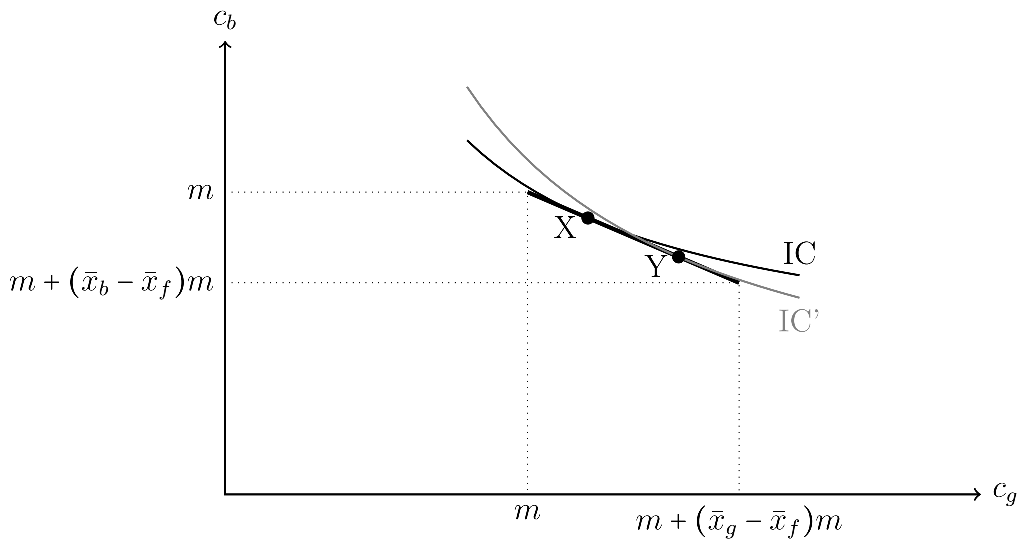

Figure 1: The bold black line segment is the set of attainable wealth combinations. IC and IC' represents the indifference curves of a salient thinker. Point X is an optimum, so the IC is tangent to the line. A relative increase in the salience of the good state increases the slope of all indifference curves. IC' indicates the new indifference curve for which tangency can be restored, taking more risk at point Y.

As illustrated in Figure 1, when $\frac{w_g}{w_b}$ rises, the indifference curve through any point (c_g, c_b) grows steeper. The optimal choice then must move towards a higher value of c_g on the budget line, weakly so at the boundaries. Since any particular c_g is obtained for the same y in any treatment, and since c_g is increasing in y, this implies that y^* is increasing in treatments. As we increase the relative decision weight on the good state, the salient thinker will bet more on the risky lottery. The following lemma expresses this as a more formal comparative statics result.

Lemma 1. *Suppose* $\frac{w'_g}{w'_b} > \frac{w_g}{w_b}$ *and* $y' > y$*. If* y' *gives weakly higher utility than* y *for a decision maker with weights* (w_g, w_b)*, then* y' *gives strictly higher utility than* y *for a decision maker with weights* (w'_g, w'_b)*.*

Lemma 1 rules out that a decision maker optimally chooses a lower y when the relative weight on the good state rises. Even if v permits non-convex indifference curves, and thus there could be multiple optimal choices for some weights, Lemma 1 implies that all optimal choices will be at least as high when the relative weight on the good state rises. Nevertheless, a few features remain worth remarking. Optimal choices sitting at one end of the bounded budget line remain there when the relative weights change would push the choice out of the line. Once choice $y = m$ is optimal, it thus remains optimal when the relative weight on the good state is higher. Symmetrically for $y = 0$. Similarly, if v has an indifference curve with a kink at an interior optimum, the optimal y may remain constant as the ratio of decision weights changes.

2.4 Comparative Statics in Salience

Lemma 1 expresses that the endogenous y is weakly increasing in the relative decision weight on the good state, $\frac{w_g}{w_b}$. In order to predict a treatment effect, we are led to the next question. Is this ratio rising or falling with treatment level Δ?

It would be particularly easy to answer this question, if a change in Δ would move salience levels $w_g = \sigma(\frac{\bar{x}_g}{\bar{x}_f})$ and $w_b = \sigma(\frac{\bar{x}_f}{\bar{x}_b})$ in opposite directions. Unfortunately, this is impossible. An implication of a constant budget line is that its slope is held constant as treatment varies. This constant slope is the ratio of $x_f - x_b$ to $x_g - x_f$. Normalizing both terms by x_f, the slope is also the ratio of $(1 - \frac{\bar{x}_b}{\bar{x}_f})$ to $(\frac{\bar{x}_g}{\bar{x}_f} - 1)$. Increasing the salience of the bad (good) state is equivalent to raising the first (second) term. Maintaining their constant ratio, salience of the two states cannot move in opposite directions.

The direction of change comes to depend on the functional form of salience function $\sigma(r)$ where $r > 1$. For this lemma, recall the definition of the upper bound on the treatment variable, $\Delta^U = \frac{\bar{e}-m}{m}$. We can conveniently define $x_s^U = \bar{x}_s + \Delta^U$ for $s = b, f, g$.

Lemma 2. *(a) With Fechner's function $\sigma(r) = K|\ln(r)|$, ratio $\frac{w_g}{w_b}$ is increasing in treatment variable Δ.*
(b) With Bordalo et al.'s function $\sigma(r) = \frac{r-1}{1+r}$, $\frac{w_g}{w_b}$ is increasing in Δ.
(c) With salience function $\sigma(r) = (r-1)^\gamma$ for $r \geq 1$ and $\gamma > 0$, $\frac{w_g}{w_b}$ is increasing in Δ.
(d) If x_f^U exceeds the harmonic mean of x_b^U and x_g^U, then with salience function $\sigma(r) = r^\gamma$ for $r \geq 1$ and $\gamma > 0$, $\frac{w_g}{w_b}$ is increasing in Δ.
(e) If x_f^U is less than the harmonic mean of x_b^U and x_g^U, then with salience function $\sigma(r) = r^\gamma$ for $r \geq 1$ and $\gamma > 0$, $\frac{w_g}{w_b}$ is locally decreasing in Δ near Δ^U.

The last case (e) of Lemma 2 serves to illustrate the general theoretical observation that monotonicity really depends on the functional form of σ which maps payoff ratios r into decision weights w. When we specify the parameters of our actual experiment below, we are going to satisfy the sufficient condition of (d). While relatively little is known about the true shape of function σ, guided by the leading suggestions of the literature in parts (a) and (b), we will hypothesize that the effect of the treatment in our experiment will be to increase the relative weight on the good state. Combined with Lemma 1, higher treatment variable Δ then raises the choice variable y.

In the experiment, we will fix parameters such that also rank-dependent Salience Theory predicts an effect. We assume $\frac{\bar{x}_f}{\bar{x}_b} > \frac{\bar{x}_g}{\bar{x}_f}$ so the bad state is most salient when $\Delta = 0$. We also assume $\bar{x}_g - \bar{x}_f > \bar{x}_f - \bar{x}_b$.[22] Then the payoff ratio in the bad state crosses that of the good state at

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

We will ensure below that $\Delta^0 < \Delta^U$. Then the bad state is most salient when $\Delta < \Delta^0$, changing to the good state being most salient as Δ passes beyond Δ^0. Rank-dependent Salience Theory predicts that y is greater once $\Delta > \Delta^0$.

In sum, we have derived the following theoretical proposition.

Proposition 1. *(i) The set of available prospects is constant to treatment.*
(ii) The salient thinker's optimal choice responds to treatment.
(iii) In the prominent specifications of continuous Salience Theory, bet amount y rises with treatment variable Δ.
(iv) In rank-dependent Salience Theory, bet amount y is piecewise constant in treatment level Δ, jumping up when Δ passes above Δ^0.

3 The Experiment

In this section, we start by describing 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 non-parametric Wilcoxon signed-rank test and a saturated Tobit regression analysis. Second, we quantify 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 mem-

[22] This assumption states that x_f is below the arithmetic mean of x_b and x_g, for any treatment level. As noted before, we will also assume the sufficient condition of (d) in Lemma 2, that x_f^U exceeds the harmonic mean of x_b^U and x_g^U. In general, the harmonic mean of two positive numbers is less than their arithmetic mean, so there is room for x_f^U to satisfy both assumptions.

bers were invited by email to participate in the experiment. In total, 473 participants finished the complete experiment and generated $1,892$ observations. Invited subjects were provided a link, a login and a password that could be used to participate. Upon logging on, they were given instructions as well as control questions. After having answered all the control questions correctly, subjects were randomly allocated to one of the 24 possible sequences of the four different treatments (i.e. decision situations) to balance our experimental design. We ended up with between 18 and 21 participants in each of the 24 different treatments. We counterbalanced the different sequences in which the four different treatments could be presented in order to eliminate the confounding of treatment and order.

Baseline payoffs were $\bar{x}_g = 1.1$, $\bar{x}_b = 0.1$, and $\bar{x}_f = 0.4$. They satisfy our assumptions that $\frac{\bar{x}_f}{\bar{x}_b} > \frac{\bar{x}_g}{\bar{x}_f}$ and $\bar{x}_g - \bar{x}_f > \bar{x}_f - \bar{x}_b$. The treatment level for equal salience is then $\Delta^0 = 0.125$. We let $\bar{e} = 160$ and $m = 100$. The probability of the good state obtaining was $p = 0.6$.

Subjects faced 4 treatment levels, labeled $T = 0, 1, 2, 3$, summarized in Table 2. The treatment values were $\Delta = 0.0, 0.1, 0.4, 0.5$. Note that the lowest two treatments are below Δ^0, while the highest two treatments are above.[23] Also, the treatments satisfy the constraint $\Delta < \frac{\bar{e}-m}{m} = 0.6$. Lastly, as already pointed out in the introduction, in order to design the experiment in such a way that the two treatments provide identical state-dependent wealth consequences, we also had to adjust the initial endowment. Changing the initial endowment across treatments might invoke the housemoney effect Thaler and Johnson (1990). The housemoney effect says that people are more willing to take risks after a recent gain. That is, if participants in our experiment perceive the original endowment as 'housemoney', they should invest less into the risky lottery going from treatment level $T = 0$ to treatment level $T = 3$. This is the opposite of our salience hypothesis presented in Proposition 1. Hence, our salience hypothesis stands in sharp contrast to the housemoney effect in the context of our experimental set-up. As a consequence, if the housemoney effect is present in our data, then our estimates of the real underlying salience effect represent a lower bound.

[Table 2 here.]

In the instructions, subjects were informed that payoffs would be expressed in actual Danish Kroner (DKK), and that only one of the four treatments would randomly be

[23] 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.

Table 2: 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

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

[Table 3 here.]

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 3 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 quartiles, it appears that 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 take two steps. First, we report results using a non-parametric Wilcoxon signed-rank test to investigate any

Table 3: 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 quartile	50	50	60	60
Median	70	80	80	85
3rd quartile	100	100	100	100
No. censored obs. at:				
Min (0)	38	33	13	17
Max (100)	172	187	202	215

treatment differences in people's investment behavior. Second, we estimate a saturated Tobit regression model.

Our hypotheses based on Proposition 1 are as follows: (i) rank-dependent Salience Theory predicts that there is at most one significantly positive jump across treatments. More specifically, parameters are chosen such that the rank-dependent version of Salience Theory predicts $b_1 = 0$ and $b_2 = b_3 > 0$. Furthermore, (ii) the appearance of more than one significant upwards jump between b_0, b_1, b_2 and b_3 is still consistent with the continuous version of Salience Theory. Remember, almost all alternative theories mentioned in the introduction prescribe $b_1 = b_2 = b_3 = 0$. [24]

[Table 4 here.]

Table 4 reports the average changes in the amount invested into the risky lottery from Treatment 0 to Treatment 1, from 1 to 2, and from 2 to 3, as well as the results from a Wilcoxon signed-rank test which non-parametrically analyzes pairwise treatment differences.

As can be seen, amounts invested significantly change from treatment level 0 to 1 as well as from treatment level 1 to 2. In contrast, the difference between treatment levels 2 and 3 is insignificant.

In addition to these non-parametric estimates, we estimate a saturated Tobit regression model (censored at 0 and 100), specification 1 of Table 5 below. It is based on the

[24] We have constrained the choice set $y \in [0, 1]$, and some observations are at the boundaries. This feature in the data may render it more likely that we fail to detect a Salience effect which is actually there. The potential bias against our hypothesis is no concern when we do detect an effect.

Table 4: Treatment Effects: Wilcoxon signed-rank test (WSR)

	Avg. change in DKK	p-value WSR
Treatment 0 to 1	4.2	0.001***
Treatment 1 to 2	5.6	0.0005***
Treatment 2 to 3	-0.8	0.90

Notes: $*$, $**$, $***$ *denote the conventional levels of significance 10%, 5% and 1% respectively*

observed amount bet by subject i at treatment level j:

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

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 $y_{i,j}$.

In the observed data, there is extraneous variation at the level of each choice. In the saturated Tobit regression model this variation is 'averaged out' by the estimation procedure. The jth-level average treatment effect (ATE-j) is embodied in the jth slope:

$$b_j = E[y_{i,j} - y_{i,0}],$$

with constant $a = E[y_{i,0}]$ being the 'control level.'

One general concern with within-subject designs is that treatment effects might be confounded by the order of presentation. As mentioned before, to eliminate the fact that carry-over effects contaminate the estimated treatment effects, we counterbalanced our experimental design such that these carry-over effects cancel out on average. We additionally report results in which we control for sequence fixed effects (specification 2) by including dummy variables for each sequence (minus 1) to the Tobit regression. The omitted reference sequence 1 presented the treatment levels in the same order as Table 2, i.e. 0, 1, 2, 3.

A further concern may be that unobserved errors are correlated across treatments at the subject level. In 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

Table 5: Average Treatment Effects: Tobit Regression Estimates

	(1)	(2)	(3)
Treatment level 1	6.18	6.20	6.18
	(2.50)	(2.49)	(3.46)
Treatment level 2	14.96	15.061	14.96
	(2.61)	(2.60)	(3.48)
Treatment level 3	15.08	15.062	15.08
	(2.82)	(2.81)	(3.49)
Constant	77.68	62.78	77.68
	(2.63)	(7.24)	(2.45)
Robust clustered errors	Yes	Yes	No
Sequence fixed effects	No	Yes	No
Sample size	1892	1892	1892
Number of clusters	473	473	-
Log likelihood	-6157.1	-6114.03	-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 estimation model adding dummies for the different sequences (minus 1). The reference sequence which is left out is sequence 1 which presented the treatment levels in the following order: 0, 1, 2, 3. Specification 3 reports plain Tobit regression estimates of the main estimation model. Standard errors are reported in parentheses beneath coefficient estimates.

valid, then there exists a clustered error problem that must be handled. We also address this concern in the following estimation of the treatment effects (specification 3).

[Table 5 here.]

Specification 1 of Table 5 estimates equation (5), controlling for within-subject correlation in the error terms, using robust clustered 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 shows that adding dummies for the different sequences to control for sequence fixed effects (over and above the carry-over effects that are controlled for by our counterbalanced experimental design) has virtually 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 from this specification are the same as in our main specification 1, but ATE-1 is now insignificant. This is likely due to the clustered standard error problem.

Furthermore, in line with our non-parametric analysis, a t-test on specification 1 indicates that ATE-1 exceeds zero ($p = 0.014$). F-statistics indicate that ATE-2 ($p = 0.0003$) and ATE-3 ($p = 0.0016$) are both significantly different from ATE-1. This implies our data is consistent with the continuous version of Salience Theory. There is a significant effect of treatment, and it goes in the intuitive direction recorded in Proposition 1 (iii).[25]

Table 5 also suggests that rank-dependent salience cannot be supported by the data. Remember, 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, as indicated in our hypothesis, rank-dependent salience distortions imply that the both effects need to be jointly true.[26]

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 attentional weight by the function $\sigma(r) = r^\gamma$ for $r \geq 1$. The salience parameter γ measures 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:

$$\ln\left(\frac{\pi}{1-\pi}\right) = \ln\left(\frac{p}{1-p}\right) + \gamma\ \ln\left(\frac{x_g/x_f}{x_f/x_b}\right). \tag{6}$$

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

[25] Recall that any significant effect, regardless of sign, is inconsistent with expected utility theory, regret theory, disappointment theory, and similarity-based theories.

[26] We are taking literally that the jump in behavior should occur at treatment level $\Delta^0 = 0.125$, and thus there should be no difference between treatment levels 0 and 0.1. The observed effect from 0 to 0.1 could be partly attributed to the prediction that the jump should occur rather close to level 0.1. This speaks to the difficulty of testing the rank-dependence model against the continuous alternative.

We need to make some assumption about the stochastic structure underlying the observations. The subject reveals a 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, to arrive at $\ln\left(\frac{v'(c_g)}{v'(c_b)}\right) + u_{i,j}$, with $u_{i,j}$ being a zero-mean normally distributed error term.

The structural estimation will be based on the optimality condition stated in equation (4). Taking into account the randomness of the marginal rate of substitution, we get:

$$\ln\left(\frac{v'(c_{b,i,j})}{v'(c_{g,i,j})}\right) = \ln\left(\frac{x_g - x_f}{x_f - x_b}\right) + \ln\left(\frac{\pi}{1-\pi}\right) + u_{i,j}. \tag{7}$$

This equation jointly depends on the decision weight π and the curvature of the value function v. It cannot be estimated without assuming some functional form for v.[27]

We consider both the constant absolute risk aversion (CARA) function, $v(c) = 1 - e^{-\rho c}$, and the constant relative risk aversion (CRRA) function, $v(c) = c^{1-\rho}/(1-\rho)$. The variation of risk aversion with respect to wealth level is not an important issue in our experiment where consequences are constant across treatments. However, by considering both functions, we can evaluate the robustness of our salience parameter estimate. We arrive at the following CARA-Tobit regression model (censored at 0 and 100):

$$y_{i,j} = a + b\ \ln\left(\frac{x_{g,j}/x_{f,j}}{x_{f,j}/x_{b,j}}\right) + u_{i,j} \tag{8}$$

with $a = \ln(3.5)/\rho$, $b = \gamma/\rho$, and $u_{i,j} \sim N(0, \sigma/\rho)$.[28] We can change regression equation (8) into a CRRA-Tobit regression model (censored at 0 and 100) by replacing y with outcome $\ln\left[\frac{c_{g,i,j}}{c_{b,i,j}}\right]$ on the left-hand side of (8).

[Table 6 here.]

Based on our regression equation (8) and its two representations, Table 6 reports Tobit regression estimates with clustered standard errors. Specification 1 of Table 6 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. We can then compute the estimates $\hat{\rho} = 0.014$ and $\hat{\gamma} = 0.39$, using the relations reported immediately after (8). Increasing the relative payoff contrast by one percent will increase

[27] By design, the term involving $(x_g - x_f)/(x_f - x_b) = 7/3$ is constant across treatments.

[28] Recognize that the CARA value function has the representation $\ln(v') = -\rho c + d$ for some d. It is then straightforward to solve for the optimal y in equation (4). Details in our Appendix.

Table 6: 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 (8) with clustered errors. Specification 2 reports the CRRA estimation model with clustered standard errors. Standard errors are reported in parentheses beneath coefficient estimates. The last two estimates are calculated by $\hat{\rho} = \ln(3.5)/\hat{a}$ and $\hat{\gamma} = \hat{\rho}\hat{b}$.

odds by 0.39 percent. Specification 2 of Table 6 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. Here, the implied parameter estimates are $\hat{\rho} = 3.06$ and $\hat{\gamma} = 0.4$. This suggests that the estimation of the salience parameter γ is robust across specifications.[29]

How do our estimates compare to estimates mentioned in the online appendix to (Bordalo et al., 2012b, p.27)? Bordalo et al. (2012b) use survey data obtained from a hypothetical study conducted on MTurk to estimate the salience parameter δ of their rank-dependent version of Salience Theory using the salience function $\sigma(r) = \frac{|r-1|}{1+r+\theta}$ and setting $\theta = 0.1$. Given this, the best fit for their data is obtained at $\delta = 0.7$. It is important to recall that our data is inconsistent with the rank-dependent version of

[29] Parameter ρ captures two different characteristics, absolute risk aversion in CARA, but relative risk aversion in CRRA.

Salience Theory. Also, our structural estimation of the parameter faces the same problem as they did, that the risk preference abruptly changes between situations with salient downside risk and situations in which the upside is salient.[30]. We can nevertheless relate their estimates to our setting and in this way better highlight the underlying relation between the estimates.

The comparison concerns the ratio of the attentional weighting functions $\frac{w_g}{w_b}$ in equation 4. In our structural estimation the ratio of the attentional weights is given by

$$\frac{w_g}{w_b} = \frac{(x_g/x_f)^\gamma}{(x_f/x_b)^\gamma} = \left(\frac{x_g x_b}{x_f^2}\right)^\gamma.$$

In Bordalo et al. (2012b)'s estimation, this ratio is instead given by

$$\frac{w_g}{w_b} = \begin{cases} \delta & \text{if } \sigma(x_f/x_b) > \sigma(x_g/x_f) \\ \frac{1}{\delta} & \text{otherwise} \end{cases}$$

with $\sigma(r) = \frac{|r-1|}{1+r+\theta}$. Using our treatment parameters, our specification of continuous Salience Theory, and the estimate $\gamma = 0.4$, this implies:

	T=0	T=1	T=2	T=3
$\frac{w_g}{w_b}$	$\left(\frac{1.1\times0.1}{0.4^2}\right)^{0.4}$	$\left(\frac{1.2\times0.2}{0.5^2}\right)^{0.4}$	$\left(\frac{1.5\times0.5}{0.8^2}\right)^{0.4}$	$\left(\frac{1.6\times0.6}{0.9^2}\right)^{0.4}$
	$= 0.86$	$= 0.98$	$= 1.07$	$= 1.07$
	Bad state salient		Good state salient	

In contrast to this, using the estimates that Bordalo et al. (2012b) obtained ($\delta = 0.7$ and $\theta = 0.1$), we get the following:

	T=0	T=1	T=2	T=3
σ_b	0.5	0.375	0.214	0.1875
σ_g	0.4375	0.389	0.333	0.269
$\frac{w_g}{w_b}$	0.7	1.429	1.429	1.429
	Bad state salient	Good state salient		

[30]They solve this issue by introducing a smoothing of the local thinker's evaluation of lotteries in the form of risk aversion in the value function

Hence, using the estimates from Bordalo et al. (2012b) implies that at treatment level T=0 the bad state is salient whereas at treatment levels T=1 to T=3 the good state is salient. With our setting this shift only occurs between T=1 and T=2 the main reason being that we assume $\theta = 0$. Relating this back to the analysis of our treatment effects, this means that there should be a significant difference between b_0 and b_1, but no difference between b_1, b_2 and b_3. Again, we observe the former but not the latter, implying that our data is more consistent with the continuous representation of Salience Theory.[31]

4 Benartzi and Thaler's Model

As already hinted at in the introduction, in principle Benartzi and Thaler's (1995) adaptation of Prospect Theory to portfolio choice can also explain that subjects allocate more to risky assets across our treatments. We now investigate this possibility in greater detail, finding that implausible parameter values would be necessary to explain our data.

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. In our experimental setting, subjects choose how to allocate m, and the ex post value of the resulting prospect is $x_s y + x_f(m - y)$. A central tenet of Prospect Theory is that people derive value from gains and losses, measured relative to some reference point. Here, by (1), the net gain is $x_s y + x_f(m - y) = c_s - c_0$, with $c_0 = e - m$ denoting the reference point.[32] Plugging in the baseline payoffs, gives $x_s y + x_f(m - y) = (\bar{x}_f + \Delta)m + (\bar{x}_s - \bar{x}_f)y$, so the value of the holding would be increasing in the treatment variable Δ 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 y, the amount bet on the risky lottery L_A, to maximize

$$V(y) = \pi(p)v(c_g - c_0) + (1 - \pi(p))v(c_b - c_0), \tag{9}$$

with π now being Prospect Theory's decision weights.

[31] An appendix to Frydman and Mormann (2017) directly assumes $\theta = .1$ and would hence make the same prediction here as Bordalo et al. (2012b). Their estimate for δ is 0.89 on average.

[32] As remarked in Section 2.2, Prospect Theory differs from Salience Theory in this focus on gains and losses.

According to Prospect Theory, the valuation function v is concave on the positive domain. The subject will bet an amount up to the point where $V'(y) = 0$. By equation (9), we therefore have that

$$\frac{x_f - x_b}{x_g - x_f} = \frac{\pi(p)}{1 - \pi(p)} \frac{v'(c_g - c_0)}{v'(c_b - c_0)}. \tag{10}$$

As the treatment variable Δ increases, the value of the holding will also increase. By the concavity of the value function, $\frac{v'(c_g - c_0)}{v'(c_b - c_0)} < 1$ and converging towards 1 as the treatment variable Δ is increased. Increasing Δ will thus increase the slope of the strictly convex indifference curve. Equality in equation (10) is then restored by betting more, just as illustrated in the earlier Figure 1.

To assess the extent to which Benartzi and Thaler's model actually explains our observations, we modify our structural model (7):

$$\ln\left(\frac{v'(c_{b,i,j} - c_{0,j})}{v'(c_{g,i,j} - c_{0,j})}\right) = \ln\left(\frac{x_g - x_f}{x_f - x_b} \frac{\pi(p)}{1 - \pi(p)}\right) + u_{i,j}. \tag{11}$$

In our estimation we will consider the power functional $v(c_s - c_0) = (c_s - c_0)^\alpha$ with $0 < \alpha < 1$, which is customarily used in applications of Prospect Theory when $c_s - c_0$ is positive. For observable outcome $z_{i,j} = \ln\left[\frac{c_{g,i,j} - c_{0,j}}{c_{b,i,j} - c_{0,j}}\right]$, we thus apply the following Tobit model with variable censoring across treatments:

$$z_{i,j} = a + u_{i,j} \text{ with } a = \frac{1}{1 - \alpha} \ln\left(3.5 \frac{\pi(0.6)}{1 - \pi(0.6)}\right). \tag{12}$$

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

$$\hat{\pi}(0.6) = \frac{e^{1.399(1-\alpha)} \frac{0.3}{0.7}}{1 + e^{1.399(1-\alpha)} \frac{0.3}{0.7}}. \tag{13}$$

There are two ways to demonstrate that the estimated parameters are implausible. First, by applying a feasible estimated range for the α-parameter in the value function v, $\hat{\alpha} \in [0.7, 0.9]$ (Tversky and Kahneman, 1992; Abdellaoui, 2000; Harrison and Rutström, 2009), we find from (13) that $\hat{\pi}(0.6) \in [0.3, 0.4]$. But these studies all jointly estimate the α-parameter and the weighting function

$$\pi(p) = \frac{p^\beta}{(p^\beta + (1 - p)^\beta)^{1/\beta}}, \tag{14}$$

and actually find that $\hat{\beta} \in [0.6, 0.9]$. Now, (14) 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.

Alternatively, assuming a more realistic underweighting $\hat{\pi}(0.6) = 0.5$ would by equation (13) imply that $\hat{\alpha} = 0.3$. 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 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. By implication, salient thinkers will put disproportional weight on salient consequences when considering the value of prospects. In particular, we observe that subjects 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 increase. 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.[33]

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

[33] It is beyond the scope of our experimental study, but it would be interesting to characterize the set of salience functions σ that are consistent with all available evidence.

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 washes out some of 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 as our structural estimation of the salience parameter measures a 'deep psychological constant.' The latter concern suggests a line along which to pursue further experimental work.

Lastly, we deliberately designed the experiment in a way to test for salience effects and control for a very large set of alternative theories which would otherwise naturally be used to predict behavior in our experimental setting. In our test of the salience effect, we took these alternative theories at face value without making any attempt to reconcile them with the data. Clearly, if we had built in, for example, a notion of 'narrow framing' similar to the one used in Benartzi and Thaler's (1995) version of Prospect Theory in one of the alternative theories mentioned in the introduction, we might have been able to obtain a hypothesis in line with our salience hypothesis. Whether such ad hoc changes to the existing theories would lead to better and descriptively more accurate representations of behavior in our setting is very interesting, but goes beyond the scope of our focus and is hence left for future research.

References

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

Allais, M., 1953. L'extension des théories de l'équilibre économique général et du rendement social au cas du risque. Econometrica 21, 269–290.

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

Bell, D.E., 1982. Regret in decision making under uncertainty. Operations Research 30, 961–981.

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

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

Booth, A., Nolen, P., 2012. Salience, risky choices and gender. Economics Letters 117, 517–520.

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

Bordalo, P., Gennaioli, N., Shleifer, A., 2012b. Salience theory of choice under risk. Quarterly Journal of Economics 127, 1243–1285.

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

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

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

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

Bordalo, P., Gennaioli, N., Shleifer, A., 2017. Memory, Attention, and Choice. Technical Report 23256. National Bureau of Economic Research.

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, 626–653.

Dertwinkel-Kalt, M., Köster, M., 2017. Salient compromises in the newsvendor game. Journal of Economic Behavior & Organization 141, 301–315.

Dertwinkel-Kalt, M., Köster, M., 2019. Salience and skewness preferences. forthcoming, Journal of the European Economic Association .

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

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

Fechner, G.T., 1860. Elemente der Psychophysik. Breitkopf und Härtel, Leipzig.

Fishburn, P.C., 1983. Transitive measurable utility. Journal of Economic Theory 31, 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 59, 667–686.

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

Hong, C.S., 1983. A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the allais paradox. Econometrica 51, 1065–1092.

Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. Econometrica 47, 263–291.

Königsheim, C., Lukas, M., Nöth, M., 2019. Salience theory: Calibration and heterogeneity in probability distortion. Journal of Economic Behavior and Organization 157, 477–495.

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

Kőszegi, B., Rabin, M., 2006. A model of reference-dependent preferences. Quarterly Journal of Economics 121, 1133–1165.

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

Kőszegi, B., Szeidl, A., 2013. A model of focusing in economic choice. Quarterly Journal of Economics 128, 53–104.

Lanzani, G., 2020. Correlation made simple: Applications to salience and regret theory.

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

Lian, C., Ma, Y., Wang, C., 2019. Low interest rates and risk-taking: Evidence from individual investment decisions. Review of Financial Studies 32, 2107–2148.

Loomes, G., Sugden, R., 1982. Regret theory: An alternative theory of rational choice under uncertainty. Economic Journal 92, 805–824.

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

Machina, M.J., 1982. "Expected utility" analysis without the independence axiom. Econometrica 50, 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, 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, 1423–1452.

Thaler, R.H., Johnson, E.J., 1990. Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. Management science 36, 643–660.

Tversky, A., Kahneman, D., 1992. Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and Uncertainty 5, 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. American Economic Review 102, 41–46.

Appendix A: Proofs

Proof of Lemma 1

Let (c_g, c_b) and (c'_g, c'_b) denote the wealth consequence pairs corresponding to choices y and y', respectively. When $y' > y$, we have $c_g < c'_g$ and $c_b > c'_b$. At weight pair (w_g, w_b), choice y' is weakly better than y if and only if

$$pw_g v\left(c'_g\right) + (1-p)\, w_b v\left(c'_b\right) \geq pw_g v\left(c_g\right) + (1-p)\, w_b v\left(c_b\right)$$

i.e.,

$$pw_g \left[v\left(c'_g\right) - v\left(c_g\right)\right] \geq (1-p)\, w_b \left[v\left(c_b\right) - v\left(c'_b\right)\right].$$

Since v is increasing, factors are positive on both sides of this inequality. When $\frac{w'_g}{w'_b} > \frac{w_g}{w_b}$, we get by implication that

$$pw'_g \left[v\left(c'_g\right) - v\left(c_g\right)\right] > (1-p)\, w'_b \left[v\left(c_b\right) - v\left(c'_b\right)\right].$$

Hence, with decision weights (w'_g, w'_b), y' is strictly preferred to y.

Proof of Lemma 2

(a). In this case,

$$\frac{w_g}{w_b} = \frac{\ln\left(x_g + \Delta\right) - \ln\left(x_f + \Delta\right)}{\ln\left(x_f + \Delta\right) - \ln\left(x_b + \Delta\right)}.$$

Its derivative with respect to Δ is seen to be positive if and only if

$$\frac{\ln\left(x_g + \Delta\right) - \ln\left(x_f + \Delta\right)}{\ln\left(x_f + \Delta\right) - \ln\left(x_b + \Delta\right)} > \frac{\frac{1}{x_f+\Delta} - \frac{1}{x_g+\Delta}}{\frac{1}{x_b+\Delta} - \frac{1}{x_f+\Delta}} = \frac{1 - \frac{x_f+\Delta}{x_g+\Delta}}{\frac{x_f+\Delta}{x_b+\Delta} - 1}.$$

This inequality is true since $\ln(r) \leq r - 1$ with strict inequality for all positive $r \neq 1$.

(b). Note that the salience of the good consequence is

$$w_g = \frac{\left(\bar{x}_g + \Delta\right) - \left(\bar{x}_f + \Delta\right)}{\left(\bar{x}_g + \Delta\right) - \left(\bar{x}_f + \Delta\right)} = \frac{\bar{x}_g - \bar{x}_f}{\bar{x}_g + \bar{x}_f + 2\Delta}.$$

Using a similar expression for w_b, we find that

$$\frac{w_g}{w_b} = \frac{(\bar{x}_g - \bar{x}_f)(\bar{x}_f + \bar{x}_b + 2\Delta)}{(\bar{x}_f - \bar{x}_b)(\bar{x}_g + \bar{x}_f + 2\Delta)}.$$

Since $\bar{x}_b < \bar{x}_f < \bar{x}_g$, this is an increasing function of Δ.

(c). Observe that

$$\frac{\bar{x}_g + \Delta}{\bar{x}_f + \Delta} - 1 = \frac{\bar{x}_g - \bar{x}_f}{\bar{x}_f + \Delta}.$$

Using a similar expression for $\frac{\bar{x}_f + \Delta}{\bar{x}_b + \Delta} - 1$, we find

$$\frac{w_g}{w_b} = \left[\frac{(\bar{x}_g - \bar{x}_f)(\bar{x}_b + \Delta)}{(\bar{x}_f - \bar{x}_b)(\bar{x}_f + \Delta)}\right]^{\gamma}.$$

Since $\bar{x}_b < \bar{x}_f < \bar{x}_g$, this is an increasing function of Δ.

(d) and (e). Since $\gamma > 0$, and since the logarithm is strictly increasing, $\frac{w_g}{w_b}$ rises if and only if

$$\ln(\bar{x}_g + \Delta) - \ln(\bar{x}_f + \Delta) - \ln(\bar{x}_f + \Delta) + \ln(\bar{x}_b + \Delta)$$

rises. Its derivative with respect to Δ is

$$\frac{1}{x_g} - \frac{2}{x_f} + \frac{1}{x_b}.$$

This derivative is positive if and only if x_f exceeds the harmonic mean of x_b and x_g, i.e.,

$$x_f > \frac{2}{\frac{1}{x_g} + \frac{1}{x_b}}. \tag{15}$$

The remaining issue is when this property holds for all permissible values of $\Delta \in [0, \Delta^U]$. The left-hand side of (15), $x_f = \bar{x}_f + \Delta$ is increasing one for one with Δ. The right-hand side of (15) always rises faster than that, for its slope is

$$\frac{2\left(\frac{1}{x_g^2} + \frac{1}{x_b^2}\right)}{\left(\frac{1}{x_g} + \frac{1}{x_b}\right)^2} = 1 + \frac{\frac{1}{x_g^2} + \frac{1}{x_b^2} - 2\frac{1}{x_g}\frac{1}{x_b}}{\left(\frac{1}{x_g} + \frac{1}{x_b}\right)^2} > 1.$$

The desired inequality (15) therefore holds for all Δ if and only if it holds at the upper end, Δ^U.

Derivation of Equation (8)

The CARA function has $v'(c) = \rho e^{-\rho c}$, so $\ln(v'(c)) = \ln(\rho) - \rho c$. The left-hand side of (7) is then $\rho(c_g - c_b)$. Substituting the definition for c_s from (1), this left-hand side reduces to $\rho(x_g - x_b)y = \rho y$, as all four treatment levels have $x_g - x_b = 1$ (Table 2). We will then arrive at (8) after dividing both sides of (7) by ρ.

On the right-hand side of (7), Table 2 shows that $(x_g - x_f)/(x_f - x_b)$=7/3 in all four treatment levels. According to (6), the second term on the right-hand side of (7) in part provides the constant level $\ln(p/(1-p))$. Here, $p/(1-p) = 3/2$ since $p = 3/5$ (Table 2). Since $\ln(7/3) + \ln(3/2) = \ln(7/2)$, we have explained constant a in (8).

By (6), the remaining part of the second term on the right-hand side of (7) is $\gamma \ln((x_g/x_f)/(x_f/x_b))$. This provides the term $b \ln((x_g/x_f)/(x_f/x_b))$ in (8). Finally, $u_{i,j}$ directly carries over from (7) to (8).

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 $T = 0$. The other decision situations were identical, besides the starting endowment and lottery payoffs. The endowment and payoffs in the other treatments can be found in Table 2.