

Dertwinkel-Kalt, Markus; Wenzel, Tobias

**Conference Paper**

## Attention and Endogenous Framing

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Behavioral and Experimental Economics, No. F04-V2

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Dertwinkel-Kalt, Markus; Wenzel, Tobias (2015) : Attention and Endogenous Framing, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Behavioral and Experimental Economics, No. F04-V2, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft

This Version is available at:

<https://hdl.handle.net/10419/112971>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Attention and Endogenous Framing\*

Markus Dertwinkel-Kalt<sup>1,2,†</sup> Tobias Wenzel<sup>1,2,‡</sup>

<sup>1</sup>Düsseldorf Institute for Competition Economics (DICE)

<sup>2</sup>Universität Düsseldorf

January 2015

## Abstract

This paper develops a theory of framing in an intertemporal context with risky choices. We provide a unifying account of existing theories of focusing by allowing a decision maker to choose her frame such that her attention is either drawn to salient events associated with an option or to the expected utilities an option yields in different time periods. Our key assumption is that a decision maker can choose her frame in a self-serving manner. We predict that the selected frame induces overoptimistic actions in the sense that subjects underrate risk but overrate chances and accordingly reveal overoptimistic actions. Hence, our theory can explain phenomena such as excessive harmful consumption (smoking, unhealthy diet) and risky investments (entrepreneurship, lotteries, gambling). We also apply our theory to static lotteries and find that classical phenomena of decision making under risk (such as the Common Ratio Allais paradox) can be rationalized by our model. We provide experimental evidence to support our claims.

*JEL Classification:* D03, D11, D90

*Keywords:* Focusing, Salience, Framing, Overoptimism.

---

\*We are grateful to Botond Köszegi, Nicola Gennaioli and Adam Szeidl for their helpful suggestions. We also thank participants of the workshop on “Limited Attention” in Copenhagen, seminar participants in Düsseldorf, Bath, and at the Royal Holloway College for their helpful comments. In addition we are grateful to Mats Köster for his research assistance.

<sup>†</sup>Email: dertwinkel@dice.hhu.de; Universität Düsseldorf, Universitätsstrasse 1, 40225 Düsseldorf, Germany.

<sup>‡</sup>Email: wenzel@dice.hhu.de; Universität Düsseldorf, Universitätsstrasse 1, 40225 Düsseldorf, Germany.

# 1 Introduction

Recently developed behavioral theories of focusing can account for a broad variety of puzzles in decision making as described in the empirical and experimental behavioral economics literature. At its core, focusing predicts that aspects of an alternative which are especially salient are overweighted, whereas less prominent, but possibly important aspects are underweighted. This distortion gives rise to an alternative's *focus-weighted utility*, and decision makers select among the alternatives to maximize their focus-weighted utility rather than their actual consumption utility.

As presented in Bordalo *et al.* (2012) and Köszegi and Szeidl (2013), decision situations can be framed in fundamentally different ways. Either of these models introduces a specific approach toward the framing of decision situations. Each frame induces distortions of attention and therefore biased decisions if decision makers focus on such features that are rendered especially salient in the respective frame.<sup>1</sup> According to the former approach, options are represented via the *states or events* they may yield. Given this frame, the probabilities of those states or events are overestimated for which the range of outcomes among the available alternatives is relatively large. This model provides an alternative rationale for violations of expected utility theory which can also be explained by prospect theory (Kahneman and Tversky, 1979). According to Köszegi and Szeidl (2013), options can be represented by their different *attributes* (which, for instance, could be *price, taste, and healthiness* for food items) and by the expected utilities which each of the options' attributes yield. This frame guides an individual's attention toward those attributes in which her range of choice is broader, i.e., in which the available options differ a great deal. Those attributes, which gather much attention, are overweighted insofar as the decision weights on these attributes are enhanced while less salient attributes are rather neglected. Therefore, Köszegi and Szeidl (2013) predict a *bias toward concentration* such that decision makers overvalue concentrated compared to dispersed advantages. In particular, this applies to intertemporal decision making if a time period in which an alternative yields a (dis)utility is considered to be an attribute of the respective option.

These two frames yield opposing predictions in many important decision situations as illustrated, for instance, in Köszegi and Szeidl (2013) with the example of smoking. If people trade-off expected utilities in time periods, the long-term risks of smoking may be underrated as they are small in expected terms for all future points in time. Thus, this frame may discard excessive smoking. If, however, people frame the same decision in terms of feasible events, then the severely adverse aspects of smoking (for instance, the risk of getting lung cancer) may become salient, so that people refrain from smoking.

---

<sup>1</sup>Several other theories of decision making under distorted attention have been recently developed, which are, however, less related to our approach. According to the model by Bhatia and Golman (2013), a decision maker's marginal utility in a good's characteristic is reference-dependent and increases in the attribute's level. In contrast, Cunningham (2012) and Bushong *et al.* (2014) propose models of relative thinking according to which the marginal utility of a characteristic decreases in its "referent" or in its range, respectively. Schwartzstein (2014), Gabaix (2014) and Woodford (2012) propose that scarce attention is allocated to attributes *ex ante*, either in an efficient way or guided through priors.

Also, in situations with uncertain future rewards, the two approaches yield different predictions. As an example, consider the decision to launch a new business. For this, an initial investment is required and future rewards are highly uncertain. Entrepreneurship offers the chance of a very high income, but on average, it is not profitable (Hamilton, 2000). The logic of Bordalo *et al.* (2012) concludes that entrepreneurship may be excessive if people focus their attention on the possible high rewards. To the contrary, the logic of Köszegi and Szeidl (2013) suggests that if people focused on the small expected returns and on the concentrated high initial investment, they would rather abstain from investing.

As both frames induce opposing actions in many setups where present investments (benefits) are to be traded off with potential future rewards (downsides), it is important to characterize which of the approaches provides the 'right' frame or the more plausible one in any given situation. The aim of this paper is to unify these two approaches in a single model which selects between the two possible frames in any decision situation. To do so we present a theory of intertemporal decision making in a risky environment that encompasses both previous approaches. We provide a unified framework which captures both frames as proposed by Bordalo *et al.* (2012) and Köszegi and Szeidl (2013). Indeed, we allow a decision maker's attention to be drawn either toward an option's expected utilities in time periods or, alternatively, her attention to be drawn to particularly vivid states associated with a risky option. We incorporate the core assumption that a decision maker can *decide* what frame to choose, i.e., she can decide how to work the information she holds on a given decision situation. Such a self-serving interpretation of information is consistent with psychological insights (e.g., Falk and Zimmermann, 2014; Dawson *et al.*, 2002; Balceris and Dunning, 2006). Therefore, in our model, a decision maker is hypothesized to choose both (1) an alternative and (2) her frame in which she evaluates the respective alternative in order to maximize her focus-weighted utility.

We derive our main results by comparing behavior in two different classes of decision situations. In one class people trade immediate benefits against future downside risks, and in the second class they trade immediate costs against future upside chances. This classification comprises many relevant decision situations. Among others, the first class contains decisions of "harmful consumption": An unhealthy eating habit may provide immediate pleasure, but may also cause future costs in the sense that it may trigger diabetes or increases the risk of heart attacks. The second class typically contains investment decisions such as the decision of whether or not to start an own business. Becoming an entrepreneur involves an immediate investment and the chances of rewards in the future.

As the key result, our model yields novel predictions concerning *overoptimistic actions* while neither Bordalo *et al.* (2012) nor Köszegi and Szeidl (2013) can, taken by itself, explain overoptimistic actions in situations with future downside risk *and* with future upside chance. We can rationalize overoptimistic behavior in the sense that decision makers underrate downside risk but overrate upside chances. With respect to the preceding examples this implies the following. Harmful consumption like smoking gives an immediate pleasure, but may cause serious diseases like lung cancer which may be realized in

any future period with a small probability. An individual framing the decision in terms of events overrates the incidence of getting lung cancer due to its severe negative outcome and therefore abstains from smoking. As this adverse outcome is unlikely, however, the expected smoking-induced harm in each future period in time is rather small compared to the large immediate pleasure derived from smoking. Thus, an individual frames the decision via expected utilities in time periods underrating the importance of the dispersed future risks and overrating the immediate benefits of smoking. As a consequence, she opts for the latter frame and smokes, even if it might be rational to abstain.

The same individual's attitude toward risk is fundamentally different if she decides whether to invest in a new, risky business. Here, expected payoffs may be rather small, but since the entrepreneurship offers the chance of a high reward, this is especially salient. Thus, the agent decides in favor of the representation via events and thereby overrates her winning chances. This can explain excessive entrepreneurship. Taken together, our model can explain overoptimistic actions in situations with future downside risk and with future upside chances in one unified framework — which existing contributions cannot.

However, such overoptimism is not universal, but has plausible limitations. A risk-averse agent, for example, will always prefer safe options (ensuring safe and positive outcomes in all time periods) over symmetric mean preserving spreads. Therefore, our model does not contradict risk-averse behavior in general. In particular, the results we produce cannot be derived by a model of risk-seeking.

The main assumption in our model, as outlined above, has been that a decision maker can freely choose her representation of a decision situation in a self-serving manner. Of course, in many instances the frame may be predefined through exogenous information or events. For example, regarding insurance against earthquakes or other natural disasters, witnessing the damages caused by the natural disasters may evoke the representation in terms of states. Therefore, an individual is forced to focus on states, in this case on the downside state (and to somewhat ignore the probability of such events and hence the expected damages). To account for such effects, we develop a version of our model where frames are preassigned and where re-framing is associated with a switching cost. In the earthquake example, avoiding thinking about the consequences (the states) of an earthquake may cause some psychological costs. In principle, these switching costs may be infinitely high.

This framework with switching costs can be used to analyze how exogenous actions or events may influence risk perception and behavior. For instance, this framework can provide a rationale for the observation that the demand for earthquake insurance increases immediately after an earthquake has happened. Witnessing the earthquake may cause a switch of frames toward a focus on states, or it may at least increase the cost of avoiding thinking of states. This may induce the decision maker to buy an insurance. The same logic can explain why the marketing activities of insurance companies should highlight the downside states, for instance, by employing salespeople who vividly describe what could happen if one does not have an insurance. But our model could also be applied to

anti-smoking campaigns by governments or other organizations. According to our approach, for instance, showing smoker lungs on cigarette packs or similar campaigns may direct attention to adverse states and be effective in reducing smoking.

To compare our results more closely to Bordalo *et al.* (2012), we also present a one-period version of our model. We ran surveys to test our model's predictions against the ones in Bordalo *et al.* (2012). The survey comprises questions which relate to different versions of the Allais Paradox (Common Ratio Allais Paradox and Common Consequence Allais Paradox). Furthermore, we also include lotteries that combine salience and risk attitudes, which are also studied in Booth and Nolen (2013). By testing for all three paradoxes both with positive (as in Bordalo *et al.*, 2012) and with negative outcomes, the study comprised six tasks. While some of our model's predictions are in line with Bordalo *et al.* (2012), there are a couple of deviations. However, our model correctly predicts the modal choice patterns in five of the six tasks, so it can explain observed patterns that Bordalo *et al.* (2012) cannot account for.

## 2 Two types of focusing

This section presents a theory of intertemporal decision making in a risky environment. In such an environment there are two tensions in the decision process of a decision maker (she). On the one hand she has to decide how to compare payoffs received in different periods, and on the other hand she has to assess risk in any given future period. The recent behavioral literature proposes two different frames through which she processes the decision problem. Our theory endogenizes the frame in which the decision situation is represented and therefore has two key ingredients.

First, a decision maker does not evaluate options according to consumption utility, but rather her attention is attracted by those aspects which the respective frame renders particularly salient. Accordingly, she puts more weight on such salient aspects. As each frame induces her to focus on specific aspects, but to neglect others, each frame gives rise to a distorted *focus-weighted utility* representation of the available options. Essentially, the frame proposed by Köszegi and Szeidl (2013) represents options via the expected utilities associated with an option at each point in time, whereas the frame proposed by Bordalo *et al.* (2012) frames the alternatives via the different states each risky option gives rise to. While the first frame induces a decision maker to overvalue expected outcomes at those points in time which are particularly salient, the second frame guides her attention to especially vivid states which therefore receive a larger weight in the decision process.

Second, our theory employs a self-serving assumption according to which the decision maker can choose the frame in which she represents the decision problem. Formally, we allow the decision maker to choose her frame (an option's expected utilities at points in time/ an option's states) as to maximize her focus-weighted utility.

## 2.1 The model

An intertemporal decision problem is uniquely determined by the following features. The time horizon  $\mathcal{T}$  determines which time periods are involved in the decision context. A decision maker chooses one option from a choice set  $\mathfrak{C}$ , while she may evaluate the different options in various frames  $x \in X$ . Each of the options yields (risky) outcomes in each of the time periods.

Precisely, for  $t < T$  with  $t, T \in \mathbb{N}$ , define the *time-horizon*  $\mathcal{T} := \{t, t + 1, \dots, T\} \in \mathbb{N}^{T-t+1}$  as the set of periods under consideration.

Let  $I$  be a finite set. For each  $\tau \in \mathcal{T}$  and  $i \in I$  there is a real-valued random variable  $C_\tau^i$  with finite support. Denote  $S_\tau$  the sample space (also called the *state space*) which is generated by the random variables  $\{C_\tau^i | i \in I\}$  and denote  $\mathcal{F}_\tau$  the corresponding canonical  $\sigma$ -algebra. We consider a probability space  $(S_\tau, \mathcal{F}_\tau, p_\tau)$  for some probability measure  $p_\tau$ , which assigns each *state of the world*  $s \in S_\tau$  its objective probability  $p_{\tau,s} := p_\tau(s)$ , such that  $\sum_{s \in S_\tau} p_{\tau,s} = 1$  holds. Note that each state  $s \in S_\tau$  can be written as a tuple which assigns each alternative  $C^i \in \mathfrak{C}$  an outcome  $c_{\tau,s}^i := C_\tau^i(s)$ , i.e.,  $s = (c_{\tau,s}^i)_{C^i \in \mathfrak{C}}$ .<sup>2</sup> A *choice option* is a vector  $C^i = (C_\tau^i)_{\tau \in \mathcal{T}}$ .<sup>3</sup> The *choice set*  $\mathfrak{C} := \{C^i | i \in I\}$  is the set of all available choice options.

We refer to  $c_{\tau,s}^i$  as the *consumption level* which choice  $C^i$  provides in period  $\tau$  and state  $s$ . A decision maker knows the probability distributions  $p_\tau$  and has to choose in period  $t$  one option from the choice set  $\mathfrak{C}$  before uncertainty is resolved and states for all periods are realized.<sup>4</sup>

Finally, a decision situation is always represented in a *frame*  $x \in X$ , where  $X$  denotes the set of available frames. A frame denotes a decision maker's mental representation of a decision situation. In the following, we will analyze two specific frames which have been proposed in the behavioral literature and which direct the decision maker's attention toward different aspects of the available choice options. While we will not directly describe what a specific frame looks like, we will define how a frame modifies the decision weights the decision maker places on different features.

*To sum up, a decision problem consists of*

- (1) a time horizon  $\mathcal{T}$ ,
- (2) probability spaces  $(S_\tau, \mathcal{F}_\tau, p_\tau)$  for all periods  $\tau \in \mathcal{T}$ .
- (3) a choice set  $\mathfrak{C}$  with finitely many choice options  $C^i = (C_\tau^i)_{\tau \in \mathcal{T}}$  and
- (4) a set of frames  $X$ .

<sup>2</sup>Synonymously, we will call  $s \in S_\tau$  sometimes also a *personal event in period*  $\tau$ .

<sup>3</sup>The random variable  $C_\tau^i$  could represent a risky asset or lottery, or an immaterial consequence such as a health impact. If an agent, for instance, decides to smoke a cigarette in period  $t$ , her respective option  $C$  may involve the negative health impact  $C_\tau$  for  $\tau > t$ . Thus, by consuming a cigarette at  $t$ , she has to decide for the entire bundle  $C$  which also comprises future negative health impacts  $c_\tau$  for  $\tau > t$ .

<sup>4</sup>For our model, it is irrelevant if the states are realized simultaneously or sequentially as the only decision is made ex ante, before any uncertainty is resolved.

A consumption level  $c_{\tau,s}^i$  provides a utility to the decision maker, which is given by an *instantaneous consumption utility function*  $u_{\tau} : \mathbb{R} \rightarrow \mathbb{R}$ . We assume that this is constant over time,  $u := u_{\tau}$  and that it satisfies diminishing sensitivity with respect to zero, that is, the function  $u(\cdot)$  is concave for positive values and convex for negative values.<sup>5</sup> For notational convenience we normalize consumption utility such that  $u(0) = 0$ . In addition, the decision maker discounts future utilities according to the exponential discounting model (Samuelson, 1937) via a discount factor  $\delta \in [0, 1]$ . As we assume that the decision maker makes the decision in period  $t$ , outcomes in this period are not discounted. In line with Köszegi and Szeidl (2013), we impose additive separability between utilities in time periods for all our approaches.

**Consumption utility.** We define the consumption utility of an option  $C^i = (C_t^i, \dots, C_T^i) \in \mathfrak{C}$  to be the present value of future expected utilities, i.e.,

$$U(C^i) := \sum_{\tau=t}^T \delta^{\tau-t} E(u(C_{\tau}^i)). \quad (\text{R})$$

with

$$E(u(C_{\tau}^i)) = \sum_{s \in S_{\tau}} p_{\tau,s} u(c_{\tau,s}^i).$$

Instead of evaluating options according to consumption utility, the decision maker evaluates her options according to the frame in which they are represented. As each frame induces the decision maker to focus on certain aspects, we call the distorted utilities which we define in the following *focus-weighted utilities*. Instead of frames we also speak of *focus types*. Broadly speaking, in the context of risky prospects in an intertemporal setting the representation of a decision problem guides the decision maker's attention either toward salient states or toward salient point in times.<sup>6</sup>

In our theory, as is detailed below, we endogenize a decision maker's representation of a choice problem by assuming that she can choose her frame. To proceed, we first need the notion of a *focusing function*.

**Definition 1** A focusing function is a continuous function  $g : \mathbb{R}^{\geq 0} \rightarrow \mathbb{R}^{\geq 0}$ .

In the two approaches which we present in the following, the focusing function assigns

---

<sup>5</sup>This property is used only in Section 7. In the remaining sections, also, for instance, linear specifications work. The property of diminishing sensitivity has been established by prospect theory (Kahneman and Tversky, 1979). In contrast to prospect theory, we assume that the reference point is fixed at zero. Also Bordalo *et al.* (2012) incorporate reference point zero in their model. In the following we will show why this specification allows us to derive all major results established by focusing and salience theory. As is well known, the assumption of diminishing sensitivity implies that a decision maker is risk-averse in the positive domain and risk-loving in the negative domain if prospects are evaluated according to consumption utility. Implicitly, this assumption also means the decision maker evaluates outcomes as gains or losses and not in terms of total wealth. For an extensive experimental study supporting these hypotheses see Bateman *et al.* (1997).

<sup>6</sup>Saliency will be defined in the following. A state (time period) is the more salient the broader the range of attainable utilities in that state (period) is.



each state (time period) a weight, the so-called *focus weight*, which is a function of the range of attainable utilities in the specific state (time period). Range refers to the difference between the maximally and the minimally achievable utilities in the respective state (time period). The larger the focus weight on a state (point in time), the more *salient* is the state (point in time). We adopt the assumption by Köszegi and Szeidl (2013), according to which the focusing function is strictly monotonic increasing, i.e., a decision maker puts a higher focus weight on a state (time period) the larger the range of attainable utilities in this state (period).<sup>7</sup>

**Assumption 1** *The focusing function  $g$  is strictly monotonic increasing.*

We say that the focusing bias is the stronger the steeper the focusing function is. In the following we present the different frames according to which decision situations can be evaluated.

**Frame (E): focus on expected utilities in time periods.** *Frame (E) guides the decision maker's attention toward different periods in time, so that she puts focus weight*

$$g_\tau := g\left(\max_{C^i \in \mathfrak{C}} \delta^{\tau-t} E(u(C_\tau^i)) - \min_{C^i \in \mathfrak{C}} \delta^{\tau-t} E(u(C_\tau^i))\right). \quad (\text{E1})$$

on the expected utility she derives from option  $C^i$  in period  $\tau$ . Define the normalization factor

$$g_E := \frac{1}{T-t+1} \sum_{\tau=t}^T g_\tau.$$

The focus-weighted utility of a decision maker who frames a decision situation via expected outcomes in time periods is then defined as

$$\begin{aligned} U^E(C^i) &:= \sum_{\tau=t}^T \delta^{\tau-t} \frac{g_\tau}{g_E} E(u(C_\tau^i)). \\ &= \sum_{\tau=t}^T \delta^{\tau-t} \frac{g_\tau}{g_E} \sum_{s \in S_\tau} p_{\tau,s} u(c_{\tau,s}^i) \end{aligned} \quad (\text{E})$$

This implies that with a focus on time periods, the focus weight depends, as in Köszegi and Szeidl (2013), on the (discounted) range of utilities a decision maker may receive in a given period. Formally, the argument of the focusing function  $g$  equals the range of attainable expected utilities possible to be received by a decision maker in a certain time period. Since, by *Assumption 1*,  $g$  is strictly increasing, time periods with a larger range of possible expected utilities among the options receive more weight in the decision process than

<sup>7</sup>This assumption is also shared by the continuous approach presented in the appendix of Bordalo *et al.* (2012). Due to notational convenience, however, they assume a discrete focusing function in their main analysis. Note that the diminishing sensitivity with respect to zero, which they assume to be satisfied by the focusing function, we incorporate in the utility function. For a further comparison of approaches see also Section 7.

periods where the options under consideration deliver more balanced expected utilities. The term  $g_E$  describes a normalization factor, according to which all focus weights are divided by their average. This allows for comparability with the focus-weighted utilities derived from the following, alternative approach.

**Frame (S): focus on states.** *Frame (S) induces the decision maker to focus her attention on salient states, so that she misinterprets the probability with which a state  $s$  occurs according to a focus weight assigned by some given focusing function  $g$ . Her subjective probability for  $c_{\tau,s}$  to occur equals*

$$\tilde{p}_{\tau,s} := \frac{g_{\tau,s}}{g_{\tau,S}} p_{\tau,s},$$

where the focus weight  $g_{\tau,s}$  is defined as

$$g_{\tau,s} := g \left( \max_{C^i \in \mathfrak{C}} \delta^{\tau-t} u(c_{\tau,s}^i) - \min_{C^i \in \mathfrak{C}} \delta^{\tau-t} u(c_{\tau,s}^i) \right) \quad (\text{S1})$$

and the normalization factor

$$g_{\tau,S} := \sum_{s \in S_\tau} p_{\tau,s} g_{\tau,s}$$

ensures that  $\sum_{s \in S_\tau} \tilde{p}_{\tau,s} = 1$ .<sup>8</sup> Denote  $\tilde{C}_\tau^i$  as the random variable which realizes  $c_{\tau,s}^i$  with probability  $\tilde{p}_s$ . Accordingly, her focus-weighted utility is defined as

$$\begin{aligned} U^S(C^i) &:= U(\tilde{C}^i) = \sum_{\tau=t}^T \delta^{\tau-t} E(u(\tilde{C}_\tau^i)) \\ &= \sum_{\tau=t}^T \delta^{\tau-t} \sum_{s \in S_\tau} \frac{g_{\tau,s}}{g_{\tau,S}} p_{\tau,s} u(c_{\tau,s}^i). \end{aligned} \quad (\text{S})$$

In contrast to frame (E), this frame guides the decision maker's focus toward states, such that the focus weights  $g_{\tau,s}$  are defined on the set of states. Weights are not attached to the expected utility of an option in a given period, but rather to states  $s$  within a given period that are more salient by offering a larger range of utilities across the different options. *Assumption 1* then implies that a state receives relatively more weight in the decision process if the range of possible utilities across the options differs more.

This frame implies that, as in Bordalo *et al.* (2012), a state  $s$  where the options' payoffs are very different receives a relatively large weight in the decision process, i.e., the subjective probability  $\tilde{p}_{\tau,s}$  exceeds the objective one. That is, states with a possibly extreme outcome (yielding either an extremely positive or an extremely negative utility) receive a relatively large weight in the decision process, even if the expected utilities across the options do not differ much at that period in time and if the probability of the extreme event is rather small. In contrast, states in which all options' payoffs are relatively balanced receive less focus weight.

---

<sup>8</sup>Note that this normalization is equivalent to Bordalo *et al.* (2012).

Another crucial feature of this approach is that small probabilities are more distorted than larger probabilities (see Proposition 1, Bordalo *et al.*, 2012). This property is exploited in Section 7, where we analyze well-known paradoxes of decision making under risk.

**Selection of the frame.** Similar to other behavioral models, we assume that the decision maker chooses an item  $C^i \in \mathfrak{C}$  not to maximize her consumption utility, but a distorted utility, here her focus-weighted utility. Unlike existing contributions, however, the distortion of the utility function, i.e., in our model her focus type, is not exogenous, but is endogenous. We posit that a decision maker chooses her representation in order to achieve a higher focus-weighted utility such that we make the following key assumption:

**Assumption 2** *A decision maker chooses an item  $C^i \in \mathfrak{C}$  and the frame  $x \in \{E, S\}$  in order to maximize her focus-weighted utility.*

That is, we assume that she employs that frame (E) or (S) which yields a higher focus-weighted utility. Given the focus, a decision maker then selects the preferred option. Formally, this decision process can be represented as a two-stage decision process:

- 1) *The decision maker chooses her frame  $x \in \{E, S\}$ .*
- 2) *Given the chosen focus type  $x$ , she chooses  $C^i \in \mathfrak{C}$  in order to maximize  $U^x(C^i)$ .*

or, equivalently, the decision maker solves

$$\max_{x \in \{E, S\}} \max_{C^i \in \mathfrak{C}} U^x(C^i). \quad (1)$$

**Remarks.** Note that our solution concept resembles the notion of a *personal equilibrium* by Kőszegi and Rabin (2006), as in equilibrium the decision on the frame represents a best response to choice  $C^i$ , while option  $C^i$  represents a best response to the chosen frame. We denote the solutions to maximization problem (1) by  $(x^*, C^*)$ .

The set of frames  $X$  a decision maker could choose from could be enlarged. For instance, we could add a rational frame  $R$  according to which the decision maker evaluates options in order to maximize her consumption utility. In principle, we could also extend the frame set toward a convex set, such that the frame  $x$  corresponds to a convex combination of the “basic” types  $R, S$  and  $E$ . The induced focus-weighted utility of such a frame  $x = x(r, s, e)$  we define, in analogy to the previous definitions, as

$$U^x(C^i) := r U_t^R(C^i) + s U_t^S(C^i) + e U_t^E(C^i)$$

for parameters  $r, s, e \in [0, 1]$  such that  $r + s + e = 1$ . Then, decision makers maximize their focus-weighted utility over  $\mathfrak{C}$  and over the frame set

$$\{x(r, s, e) | r, s, e \in [0, 1], r + s + e = 1\}.$$

An equilibrium choice is denoted by  $(x^*(r, s, e), C^*)$ . Note that the optimal focus-weighted utility can always be achieved through a corner solution, i.e., either  $r = 1$ ,  $s = 1$  or  $e = 1$ . In our applications in the next section, it will never be optimal to choose frame  $R$ , so that excluding the rational considerations for the decision maker will not constrain our set of equilibria.

## 2.2 Discussion of the psychological foundations and the constituents of our model

Here, we discuss the two main psychological assumptions underlying our model, attention to salient features and self-servingness in choosing a frame and directing attention, as well as our model's crucial constituents.

**Salience.** It is well known that people's attention is often drawn toward more vivid attributes, and that those vivid attributes receive a disproportionate weight in the decision process when comparing different options (Taylor and Thompson, 1982). A well-known example comes from Schkade and Kahnemann (1998) who argue that when comparing the quality of life in California and the Midwest, a decision maker attaches a disproportionate weight to climate and weather conditions compared to other attributes where the regions are more similar.

Which features of a given decision situation attract the decision maker's attention depends on the representation of the decision problem. In decision situations under uncertainty, a frame may either direct attention toward salient states associated with an option (for instance, the large gains of winning the jackpot of a lottery) or toward the expected outcome an option yields at a certain point in time (for instance, a lottery's expected payoff). While other frames guiding the decision maker's attention toward different aspects could exist, we restrict our paper on these frames as these have an intuitive appeal and can explain important choice patterns (Bordalo *et al.*, 2012; Köszegi and Szeidl, 2013). As we will argue in the following, our model also generates insights from the closely related frames proposed by Bordalo *et al.* (2013) or Thaler (1980).

**Self-servingness.** Self-serving judgments or self-serving assessments of information represent also a robust psychological mechanism which can be observed in many circumstances. For example, people's judgment of what is fair is influenced by self-interest (Messick and Sentis, 1979), but also objective information is assessed in self-serving manners (Dawson *et al.*, 2002; Balcetis and Dunning, 2006). Whereas according to Bordalo *et al.* (2012) and Köszegi and Szeidl (2013) attention is fully shaped by the environment, an experiment by Falk and Zimmermann (2014) finds that this is not entirely true and conclude that "subjects can actively manage attention in a self-serving way".

In our model, self-servingness is reflected by positing that a decision maker's assessment of risky situations is influenced by a self-serving motive: the decision maker may

decide on her frame in order to maximize her focus-weighted utility. As will be seen later, the assumption that the frame is chosen is the key assumption of our model, and helps us explain why decision makers may assess risk quite differently when faced with large upside or downside risks.

Our approach is related to models of optimal expectations (Brunnermeier and Parker, 2005) and wishful thinking (Mayraz, 2011). We will compare these models with our approach in Section 6.

**Constituents of the model.** In line with the two approaches which our model builds on, we assume that certain features of our model are exogenous and uniquely specified for a given decision situation, i.e., (i) the choice set  $\mathcal{C}$ , (ii) the state spaces  $S_\tau$  and (iii) the attribute space. Therefore, first, we treat the options a decision maker considers before making a choice as fixed. As we will focus on binary decisions of whether to undertake or abstain from an action, the assumption that the choice set is uniquely defined seems natural. However, choice set effects could be very important for the attention-models we build on as well as for our model. Additional items in a decision maker's choice set could render different aspects salient and therefore trigger decoy- or attraction effects (see in particular Bordalo *et al.*, 2013), which, however, are not the focus of the present paper. Therefore, the assumption of an exogenous choice set does not represent a severe limitation for our current model. Note, however, that in general we could extend our analysis toward *consideration sets*, which may differ from the actual choice sets if some available options are not considered by the decision maker or if unavailable options come to mind, for instance, because they have been consumed in the past.

Second, we treat the state space and the distributions of states as fully specified and as common knowledge. In particular for our model's application to lotteries and experiments (see Section 7), this assumption is plausible and does not represent a limitation of our approach (see also Bordalo *et al.*, 2012). In our other applications, however, we have to restrict our analysis to tractable state spaces and mostly binary choices and argue why these simplifications provide generalizable insights. In addition, the probability distribution of the states is well defined and is common knowledge (see also Bordalo *et al.*, 2012). This implies that people have rational expectations insofar as they know how their actions influence their future outcomes, which is in line with Köszegi and Szeidl (2013). More critical is our assumption of separate state spaces for each period, while an alternative configuration would define one state space for the entire horizon. As the latter, however, is untractable, we impose this simplifying assumption.

Third, attributes are exogenous and are defined by the choice context (for a discussion of this issue, see Köszegi and Szeidl, 2013). As we focus on intertemporal decisions, our model is based on the assumption that utilities obtained in different time periods represent different attributes. Also all other attention-based theories our model is related to treat the set of attributes also as an exogenous constituent of their model, such that an endogenization of the attribute space is a promising area for future research.

In addition we assume a “pain of paying,” i.e., an immediate utility from monetary transactions. For a thorough discussion of this assumption see Köszegi and Szeidl (2013). In particular, we need this assumption in the next section when we analyze investment decisions into lottery tickets or risky businesses.

### 2.3 On the relation to Bordalo *et al.* (2012) and Köszegi and Szeidl (2013)

Frame (E) is formally equivalent to the model by Köszegi and Szeidl (2013) as all our basic model assumptions, besides the assumption on the curvature of  $u(\cdot)$  and the normalization factor, are imposed by Köszegi and Szeidl (2013). The property of diminishing sensitivity which  $u(\cdot)$  is assumed to satisfy with respect to zero represents a refinement of the model by Köszegi and Szeidl (2013), which the authors also propose and discuss in their Section III.D., but do not incorporate in their main body. We incorporate this property as it is a crucial component of the model by Bordalo *et al.* (2012). Whereas they incorporate this property for the focusing, but not for the utility function, in Section 7 we derive why these two approaches are equivalent in most regards. The normalization factor is irrelevant for the model by Köszegi and Szeidl (2013), however, it is important in order to compare the focus-weighted utilities derived from either frame. The normalization ensures that the rational valuation is obtained if no point in time is particularly salient, i.e., if at no point in time the range of attainable utilities is larger than at another point in time.

While Köszegi and Szeidl (2013) assume that all subjects do not have a bias of risk processing, but have only a distorted perception of expected utilities, our notation incorporates the explicit underlying state space in order to also account for risk biases as captured by frame (S). This frame mirrors the intuition by Bordalo *et al.* (2012) that the subjective probabilities of those states, in which the range of utility outcomes is large, are magnified. As the normalization of the probability distortions is taken from Bordalo *et al.* (2012), frame (S) mirrors their continuous model.<sup>9</sup> The normalization ensures that if no state is particularly salient, i.e., if the range of utility outcomes is the same for all states, then frame (S) induces the rational valuation. In addition, this normalization ensures that low-probability states are subject to stronger distortions through salience than high-probability states, which is proven in Bordalo *et al.* (2012)’s Proposition I and which becomes important in our Section 7.<sup>10</sup>

Therefore, frames (E) and (S) mirror the approaches by Bordalo *et al.* (2012) and Köszegi and Szeidl (2013) with a unified notation. Our endogenization of the decision maker’s frame via our assumption of self-servingness assigns each decision situation a unique representation, i.e., frame (E) or frame (S), and induces a self-selection of the decision

<sup>9</sup>Whereas in the main body the authors assume rank-based salience, they also suggest continuous salience weights which we map in our setup.

<sup>10</sup>Extending the model by Bordalo *et al.* (2012) toward choice sets with more than two options reveals one further difference to Köszegi and Szeidl: according to the former, focus-weights are option-specific, while they are not option-specific in the latter model. As we, however, analyze only binary choices, this distinction vanishes.

maker into one of the two models.<sup>11</sup>

## 2.4 Plan of the paper

In the next section (Section 3) we apply this general model of intertemporal decision making by considering two polar cases of focusing on binary decisions: A decision maker either trades off an immediate benefit against a future downside risk or she trades off an immediate cost against a future upside risk. Many real-life situations are covered by those two polar cases. In Section 4 we introduce a version with switching costs to account for situations where a decision maker is limited in the choice of a focus type. Section 5 discusses several extensions to our baseline model, while in Section 6 we compare our model to other theories of manipulable expectations. In Section 7 we abstract from intertemporal decisions by applying the model to a one-period setup where the decision maker faces the choice between several lotteries.

## 3 Overoptimistic actions

Our model makes distinct predictions concerning intertemporal decision making, which we will explore in this section. Each intertemporal decision trades relative benefits in some time periods for relative costs in other time periods. We distinguish between two major, opposing classes of decision situations: people may either trade *immediate benefits for future costs* or *immediate costs for future expected rewards*.

Any kind of harmful consumption typically belongs to the first class: people smoke, drink alcohol or engage in unhealthy eating habits. In other words, people take risks and potential future costs in terms of bad health or a reduced life expectation to enjoy immediate pleasure. By smoking now, people have a higher risk of getting respiratory diseases or lung cancer. By drinking too much alcohol, people risk getting liver diseases and hepatic cancer, and an unhealthy diet may trigger the development of diabetes. Whereas excessive harmful consumption typically implies a low value of life (Ippolito and Ippolito, 1984), our model gives a novel explanation for such behavior: people are overoptimistic about the consequences of their actions. Below we argue that, according to our model, in those situations decision makers choose frame (E) and thereby direct their attention to the expected utilities in different periods associated with an option. Consequently, they somewhat ignore the downsides as these are blurred over time: the enjoyable immediate effect yields a large positive utility, while the adverse components of the choice are, in each future period, small in terms of expected utility. Consequently, focusing on relatively large expected utilities in different time periods overrates the upsides, but underweights the downsides from present consumption.

---

<sup>11</sup>We do not analyze a focus on arbitrary points or states as our goal is to unify existing research on attention-based biases, for which empirical support has been collected already. Furthermore, such an approach would lose all its predictive power.

Investment decisions belong to the second class: people incur immediate costs to gain benefits in the future. A prime example is entrepreneurship where a decision maker invests in a new business, and the chances of success are highly uncertain. Existing research suggests that people are too optimistic when becoming an entrepreneur (Camerer and Lovo, 1999; Koellinger *et al.*, 2007). Failure rates are high and, on average, it is not profitable to become an entrepreneur (Hamilton, 2000; Moskowitz and Vissing-Jorgensen, 2002). Our model can explain this overoptimistic, probably excessive, entrepreneurial activity. When deciding to invest, people direct their attention on the chance of becoming "the next google" while neglecting that on average it is far more likely that the new business will result in failure.

A similar example is the phenomenon that so many people engage in games of chance, which often have a significant negative expected return. In Germany alone, 25.5 million people participated in gambling in 2011, while 11.6 million took part in the weekly Saturday lotto. Especially lotto subscriptions are widespread, according to which agents make a certain payment in order to participate in each weekly lotto over a longer period. However, only 48% of the stakes are on average returned to the gamblers (Beckert and Lutter, 2007). Our model yields an intuitive reason for engaging in unprofitable gambling: people focus on the winning states and therefore overrate their chances of winning.

**Specification.** We restrict ourselves to binary decisions as such a simplified setup suffices to generate our model's main insights.<sup>12</sup> We also abstract from any time discounting and set  $\delta = 1$ . However, our results are fully transferable to setups in which future (perceived) utilities are discounted.<sup>13</sup>

We consider the choice between a risky option  $c^u$  and a safe option  $c^c$  (which we interpret as *abstaining from the risky option*) in a period  $t$ . The safe option realizes a payoff  $c_\tau^c$  at each point in time  $\tau$ . For simplicity of presentation we normalize the payoff of the safe option to  $c_\tau^c = 0$  for all  $\tau$ . Regarding the uncertain option  $c^u$ , the two payoffs  $c_{\tau,1}^u$  and  $c_{\tau,2}^u$  are feasible. Therefore,  $S_\tau$  can be assumed to consist of two states, indexed by 1 and 2, for all  $\tau > t$ . For notational convenience, we order the states such that the utility realized by  $c^u$  decreases in the number of the state, i.e.,  $u(c_{\tau,1}^u) \geq u(c_{\tau,2}^u)$ .

### 3.1 Harmful consumption (or future downside risk)

We start with a situation where a decision maker trades off immediate (safe) benefits against future downside risks. This might include situations of harmful consumption (such as smoking or an unhealthy diet) where the gratification is immediate, but is accompanied by a severe future (downside) risk. In this case, the future risk might consist of harmful health consequences.

<sup>12</sup>We are mainly interested in analyzing whether people engage in a certain behavior, and not to which degree they engage in such actions. Hence, we consider a setup with a binary decision. Moreover, arbitrary consideration sets yield vast state spaces and, therefore, cannot be analyzed in general.

<sup>13</sup>In that case, however, without knowledge of the exact shape of the focusing function, our formulas would be cluttered as the focusing function would have to be evaluated at many points.



In our setup, such decision situations can be represented by choices from a choice set  $\mathfrak{C} = \{c^c, c^u\}$ , where  $c^c$  represents the safe option of abstaining from harmful consumption, while  $c^u$  is the potentially harmful activity that imposes a risk upon the decision maker.

We model the risky option of harmful consumption as follows. Let  $F := u(c_t^u) > 0$  denote the immediate safe consumption utility of the risky option to be obtained in period  $t$ . Since immediate gratification is assumed to come at the cost of future risk,  $c_\tau^u$  for  $\tau > t$  may yield two different outcomes. Therefore, for all periods  $\tau > t$  we distinguish between two states. Either, the risky option yields the same outcome as the safe option ( $s_1$ ), or it yields an adverse outcome ( $s_2$ ) such that  $-L := u(c_{\tau,2}^u) < 0$ . For instance, in the smoking example, state 1 represents the outcome that the smoker is not diagnosed with lung cancer, and state 2 represents the outcome where she is. We assume that the probability space is identical over all  $\tau > t$ .<sup>14</sup> Let  $p := p_{s_2}$  denote the probability of state 2 at each  $\tau > t$ . We denote the expected utility that comes with the initial choice of the harmful product in each future period  $\tau > t$  by  $f := -E(u(c_\tau^u))$ . We consider  $L > F > f$ . That is, we consider situations where the utility loss ( $-L$ ) in the downside state is relatively large but occurs with a relatively small probability.

We now analyze the choices made by a decision maker whose attention is drawn on states (S) or on expected utilities (E). Note that the assessment of the safe option is independent of the chosen frame since  $U(c^c) = U^S(c^c) = U^E(c^c) = 0$ .

To start with, suppose that the decision maker evaluates the risky option according to frame (S). We first note that the downside risk of the risky option is the more salient state, and hence,  $g_{\tau,s_2} = g(L) > g_{\tau,s_1} = g(0)$ . This implies that the decision maker tends to overweight the downside risk compared to a rational decision maker. That is:

$$\begin{aligned} U^S(c^u) &= F - (T - t) \left( Lp \frac{g(L)}{pg(L) + (1 - p)g(0)} \right) \\ &< F - (T - t)Lp = U(c^u). \end{aligned}$$

Due to the higher focus weight on the downside the decision maker attaches less focus-weighted utility than a rational decision maker to the risky option. In other words, focus-weighted utility falls short of consumption utility, and the decision maker with focus (S) would behave more risk-averse than a rational decision maker.

Next consider the case where the decision maker evaluates the options in frame (E). In that case, the decision maker evaluates the options according to the expected utility provided in each period. The focus-weighted utility experienced by the decision maker

<sup>14</sup>Then, it is implied that  $c_\tau^u$  for  $\tau \in \mathcal{T}$  are independently and identically distributed (*iid*) which is quite a strong assumption that might be too strict in many applications. While independence of  $\{c_\tau^u | \tau \in \mathcal{T}\}$  is a formal component of our model, we relax this assumption in Section 5.1 and show that our main results also carry over to correlated random variables. For instance, in the smoking example, a more plausible assumption would be that if the decision maker gets lung cancer at some period in some period  $t$ , then this negative health impact is also present in all periods  $\tau > t$ .

with frame (E) is given by

$$U^E(c^u) = \frac{(T-t+1)}{(g(F) + (T-t)g(f))} (g(F)F - (T-t)g(f)f). \quad (2)$$

With  $F > f$  it follows that  $g(F) > g(f)$ , and an agent puts more weight on the concentrated upside of the risky option (the immediate benefit in period  $t$ ) than on the dispersed downsides.<sup>15</sup> Since the agent puts too much weight on the immediate gratification compared to the downside, the perceived utility exceeds true consumption utility,  $U^E(c^u) > U(c^u)$ . Therefore, a decision maker tends more toward the risky option.

Indeed, with focus (E) an agent chooses the risky option as long as  $U^E(c^u) > 0$ , which holds if and only if

$$\frac{g(F)}{g(f)} > \frac{(T-t)f}{F}. \quad (3)$$

If the focusing function is sufficiently steep, i.e., the fraction  $g(F)/g(f)$  is sufficiently large, then the agent will choose  $c^u$  over  $c^c$ . However, the risky option represents the suboptimal choice,  $U(c^u) < U(c^c)$ , as long as the dispersed risks outweigh the immediate benefits,  $F < (T-t)f$ . Thus, if the focusing effect is strong enough, an agent will be overoptimistic concerning her consumption decision's future costs.

Since  $U^E(c^u) > U(c^u) > U^S(c^u)$ , the focus-weighted utility is higher when the decision maker is of focus type (E), and according to the *Assumption 2*, the decision maker would choose to be of type (E). By focusing on the case where a rational decision maker would not choose the risky option (i.e.,  $F < (T-t)f$ ), we summarize our preceding discussion:

**Proposition 1** *Suppose  $L > F > f$  and  $F < (T-t)f$ . Then, the decision maker underrates the unfavorable risk by choosing frame (E). She behaves overoptimistically and chooses the risky option if and only if (3) holds.*

The proposition provides conditions as to when and why a decision maker may decide to engage in risky options with large downside risks, even when it is not rational to do so. By choosing to focus on aggregate outcomes at each point in time, she avoids explicitly thinking of the downside state of the risky option. Therefore, the overoptimistic decision maker is less concerned regarding the potential downsides than is rational and is more likely to engage in the risky option.

Note that extending the set of frames the decision maker could choose from by adding the rational frame (R) and convex combinations of (E), (S) and (R) does not affect the choice of the decision maker. In order to maximize her focus-weighted utility, she wants to neglect downside risks and overweight upside chances. Therefore, she will abstain from rational valuations and choose either frame (E) or (S).

---

<sup>15</sup>If  $f > F$ , then both types of local thinking predict an overweighting of the downsides, such that in equilibrium the individual will never engage in harmful consumption.

Furthermore, if there is an ex ante positive risk that the adverse outcome is realized in each period  $\tau > t$  (independent from the decision maker's consumption decision), our qualitative predictions are not affected. Then, three states are feasible, in which consumption and no consumption yield both utility 0 (state 1), both utility  $-L$  (state 2), or exclusively consumption induces the adverse outcome, while no consumption does not (state 3). Here, frame (S) renders state 3 salient, while frame (E) allows the decision maker to focus on the immediate benefit derived from smoking. Therefore, a sufficiently strong focusing bias induces the decision maker to behave optimistically, to choose frame (E) and underrate consumption risks.

### 3.2 Taking bets (or future upside risk)

Now consider the opposite case where the risky option includes a (non-stochastic) immediate investment cost and the returns are random and dispersed over all future periods. For instance, an agent may decide whether or not to become an entrepreneur and invest in a new business.

As before, we normalize the utility from the safe option (i.e., not becoming an entrepreneur) as  $c_\tau^c = 0$  for all  $\tau$ . The risky option (becoming an entrepreneur) involves an investment  $-H := u(c_t^u)$  in period  $t$ , and gives rise to a future stochastic payoff. The state space is assumed to be constant across periods in time and to comprise exactly two states. We define state 1 as the loss state, in which the lottery pays 0 (the new business is a failure and the initial investment is lost), and we define state 2 to be the winning state ("becoming the next google") yielding utility  $G$  with probability  $q$ . The expected utility from the risky option in future period  $\tau$  is then  $h := E(u(c_\tau^u))$  for all  $\tau > t$ .<sup>16</sup> We consider  $G > H > h$  such that the upside of the risky option,  $G$ , is very prominent, but occurs with a relatively small probability.

In principle, our arguments from above are turned upside down if we consider future upside risks instead of a future downside risk. Suppose the decision maker uses frame (E). Since  $g(h) < g(H)$ , this implies that the one-time investment cost is given more weight relative to the expected future benefits. This also implies that  $U^E(c^u) < U^E(c^c) = 0$  holds.

Now suppose that the decision maker's attention is drawn toward potential states of the lottery so that the decision maker uses frame (S). In this case, the winning state, which yields  $G$ , is particularly salient. Due to  $g(G) > g(0)$ , the decision maker has a large focus weight on the upside of the investment lottery. As the lottery's upside is given more weight by the biased decision maker than by a rational decision maker, the focus-weighted utility derived from frame (S) exceeds the true consumption utility. In fact, with a focus on salient states, a decision maker would choose the risky option over the safe option if

---

<sup>16</sup>Note that this implies that the size of the winning state, measured by  $G$ , does not change over time. Whereas these are restricting assumption in order to simplify the state space, our analysis also holds with respect to different-sized gains. However, more assumptions about the shape of the focusing function would then have to be made.

and only if  $U^S(c^u) > U^S(c^e) = 0$ , which holds if and only if

$$-H + (T - t)G \cdot \frac{qg(G)}{qg(G) + (1 - q)g(0)} > 0, \quad (4)$$

or equivalently

$$\frac{g(G)}{g(0)} > \frac{H(1 - q)}{q((T - t)G - H)}. \quad (5)$$

If the focusing bias is strong enough, i.e., the fraction  $g(G)/g(0)$  is large enough, then an agent of focus type (S) will choose to invest even if its expected payoff is negative, i.e.,  $H > (T - t)h$ , and a rational decision maker would behave risk-averse by picking the safe option.

Since  $U^S(c^u) > U(c^u) > U^E(c^u)$ , in the case of future upside risk, the decision maker will choose frame (S) and thereby obtains a higher focus-weighted utility. However, by doing so, she *decides* to ignore expected outcomes and focus on beneficial states. Therefore, agents invest excessively in a new business as they focus on their chances and are overoptimistic about the likelihood of success. We consider a situation where it is rational not to invest in the risky option (i.e.,  $H > (T - t)h$ ) and find:<sup>17</sup>

**Proposition 2** *Suppose  $G > H > h$  and  $H > (T - t)h$ . Then, the decision maker overrates favorable risk by choosing frame (S). She behaves overoptimistically and chooses the risky option if and only if (5) holds.*

### 3.3 Attention and overoptimistic actions.

Our previous results are summarized in the following proposition. As our core predictions are overoptimistic actions, we subsequently discuss the psychological phenomenon of overoptimism.

**Proposition 3** *Fix  $L > F > f$  and  $G > H > h$  as well as  $p, q > 0$ . An agent deciding whether to take immediate benefits  $F$  for risky future expected costs  $f$  (where loss  $L$  occurs with probability  $p$ ) in all future periods chooses this option if and only if (3) holds. An agent deciding whether to take immediate costs  $H$  in order to obtain  $G$  with probability  $q$  at each future period chooses the option if and only if (5) holds. If both conditions hold, an agent underrates adverse, but overrates favorable risks and therefore engages in overoptimistic actions.*

According to our model, people reveal overoptimistic actions. If focusing effects are sufficiently strong, then a decision maker in our model tends to ignore downside risks and engage in risky behavior which a rational agent would abstain from. The very same agent, however, also overrates upside risks and therefore bears unreasonable, immediate investment costs which are unlikely to pay off.

<sup>17</sup>As in the previous paragraph, extending the frame set with respect to the rational frame does not affect the decision maker's equilibrium choices.

In our model, decision makers become risk-seeking in certain domains as a result of their susceptibility to salience and their tendency to think locally. Hereby, we provide novel intuition for the fact that agents ignore substantial downside risks *and* reveal risk-seeking behavior by investing excessively in unlikely high-stake gains.

We could transfer our insights from this section to other domains besides intertemporal decisions. Our results in the paragraph *harmful consumption* imply that people take too many risks in several dimensions in order to enjoy a certain large benefit in one dimension, whereas the paragraph *taking bets* reveals that people may sacrifice too much in one dimension to achieve unlikely high benefits in another dimension of their choice.

Our results coincide with the robust finding in psychology that people are overly or unrealistically optimistic. For example, empirical studies show that people often tend to be too optimistic about their future prospects, overrate the chances of positive events or underrate the risks of negative events (e.g., Weinstein, 1980; Taylor and Brown, 1988; Weinstein, 1989).<sup>18</sup> In particular, entrepreneurs are often much too optimistic regarding their chances of making a success of their enterprise (Koellinger *et al.*, 2007). Our model produces such overoptimistic behavior.

It is important to note that our results are fundamentally different from the findings in Bordalo *et al.* (2012) and Köszegi and Szeidl (2013). Neither of these decision models can jointly explain overoptimistic actions in situations with future downside risk *and* with future upside chances.

In Bordalo *et al.* (2012) a decision maker overvalues states where the options differ more. Therefore, in situations with large upside risk this model predicts that people would excessively decide in favor of the risky option. This coincides with our prediction. However, with large downside risk, Bordalo *et al.* (2012) predict people to overrate the salient downside so that these would be less likely to engage in the risky option. This contrasts with our predictions.

The focusing model by Köszegi and Szeidl (2013) produces the opposite predictions to Bordalo *et al.* (2012). With a focus on expected utilities in different time periods, their main result is a bias toward concentration. This implies that in situations where an agent trades off an immediate benefit against a future downside, concentrated benefits would be overweighted and actions would be more likely to be overoptimistic. In contrast, in the investment situation, the decision maker would overestimate the one-time investment costs and therefore be less likely to choose the risky action.

To sum up, our model shows how self-serving interpretations of decision situations can counteract pure attention effects and can therefore be seen as a selection criterion between Bordalo *et al.* (2012) and Köszegi and Szeidl (2013). In particular, we would argue that both theories can only partially explain observed behavior, but our unified approach with a self-serving focus can explain overoptimistic behavior with both future upside and

---

<sup>18</sup>Psychologists have shown overoptimism to be present in a wide variety of contexts. For an overview see, for instance, Shepperd *et al.* (2013). This has also been shown to be significant in situations of economic interest, such as market entry (Camerer and Lovallo, 1999) or investment decisions (Malmendier and Tate, 2005).

downside risks.

### 3.4 Frame choice and skewness

This subsection provides more intuition on our mechanism of endogenous framing by studying the effects of skewness on the frame choice. We will show that payoff distributions skewed toward positive and negative outcomes affect the chosen frames in a different manner. In particular, with outcome distributions skewed toward gains, the frame (S) is more likely to be chosen, while a skew toward losses implies that frame (E) is more likely to be selected.

Suppose that a decision maker can choose between a risky option  $c^u$  and a safe option  $c^c$  yielding a utility of zero in every period  $\tau \in \mathcal{T}$ . Assume that for periods  $\tau \in \mathcal{T}_1$ , where  $\mathcal{T}_1 \subseteq \mathcal{T}$ , the risky option consists of a lottery which yields a utility gain  $G_\tau$  with probability  $p_\tau$  and a loss of  $-L_\tau$  with probability  $1 - p_\tau$  such that  $p_\tau G_\tau - (1 - p_\tau)L_\tau = 0$ . If option  $c^u$  is evaluated with frame  $x \in \{S, E, R\}$ , we denote the sum of its focus-weighted utilities over all periods  $\mathcal{T} \setminus \mathcal{T}_1$  as  $c_x^u \in \mathbb{R}$ .

Note that the evaluation of the safe option is independent of the frame and, hence,  $U^S(c^c) = U^E(c^c) = U^R(c^c) = 0$ . Given frame (S),  $c^u$  yields the focus-weighted utility

$$U^S(c^u) = c_S^u + \sum_{\tau \in \mathcal{T}_1} \frac{p_\tau g(G_\tau)G_\tau - (1 - p_\tau)g(L_\tau)L_\tau}{p_\tau g(G_\tau) + (1 - p_\tau)g(L_\tau)}. \quad (6)$$

In the following we consider two such risky options (denoted  $A_1$  and  $A_2$ ) which are defined as follows. First, they are identical with respect to all periods  $\tau \in \mathcal{T} \setminus \mathcal{T}_1$ . Second, in all periods  $\tau \in \mathcal{T}_1$  both options are positively skewed, i.e.,  $G_\tau^1, G_\tau^2 > L_\tau$  where  $G_\tau^i$  denotes the gain option  $A_i$  provides in period  $\tau$ . Denote  $p_\tau^i$  the probability with which  $G_\tau^i$  is realized. Third, we say that *option  $A_2$  is (weakly) more positively skewed than  $A_1$*  if  $G_\tau^2 \geq G_\tau^1$  for all periods  $\tau \in \mathcal{T}_1$ .

Analogously, let  $A'_1$  and  $A'_2$  be two options which differ only with respect to periods  $\tau \in \mathcal{T}_1$ . In these periods, both options are negatively skewed, i.e.,  $L_\tau^1, L_\tau^2 > G_\tau$  where  $L_\tau^i$  denotes the loss option  $A'_i$  provides in period  $\tau$ . Furthermore, we say that *option  $A'_2$  is (weakly) more negatively skewed than  $A'_1$*  if  $L_\tau^2 \geq L_\tau^1$  for all periods  $\tau \in \mathcal{T}_1$ . The following Proposition, which we prove in the Appendix, evaluates the effect of more skewed payoff distributions on the frame choice if a risky option is compared to the safe option  $c^c$ .

**Proposition 4** *i) Let  $A_1$  and  $A_2$  be two risky options, where  $A_2$  is more positively skewed than  $A_1$ . If  $U^S(A_1) > U^E(A_1)$  holds given that the decision maker chooses from choice set  $\mathfrak{C}_1 := \{A_1, c^c\}$ , then also  $U^S(A_2) > U^E(A_2)$  given that the decision maker chooses from choice set  $\mathfrak{C}_2 := \{A_2, c^c\}$ .  
ii) Let  $A'_1$  and  $A'_2$  be two risky options, where  $A'_2$  is more negatively skewed than  $A'_1$ . If  $U^E(A'_1) > U^S(A'_1)$  for  $\mathfrak{C}'_1 := \{A'_1, c^c\}$ , then also  $U^E(A'_2) > U^S(A'_2)$  for  $\mathfrak{C}'_2 := \{A'_2, c^c\}$ .*

The Proposition shows that more skewed outcome distributions have different effects on the frame choice depending on whether the skew is in the gain or the loss domain.

Part i) shows that a more positively skewed distribution makes it more likely that decision makers adopt frame (S). To the contrary, part ii) shows that with a more negatively skewed distribution the decision maker tends to adopt the frame (E). The rational frame (R) will be adopted if the outcome distribution is symmetric as we show in Section 5.2.

## 4 Focusing, information, and switching costs

In this section, we discuss a setup where choosing a frame seems less feasible and therefore introduce a version with switching costs. The frame of a decision maker may not always depend on her choice, as is assumed in our model, but in many situations it may also depend on the amount of available information or the vividness of the different feasible states.

In particular, a decision maker's focus type may also be influenced by events or actions exogenous to the decision maker. Natural phenomena (disasters), firms' marketing activities or governments' information campaigns may affect how people assess the consequences of risky decisions. For instance, in terms of earthquake insurance decisions, the occurrence of a disastrous earthquake may render the adverse state much more vivid than before. Also, an insurance company's marketing strategy may include salespeople who highlight averse states and thereby increase their vividness. A government's information campaign, for instance an anti-smoking campaign, may include the use of images of a smoker's lungs, rendering the downside of smoking more vivid than before. We would argue that, in all those cases, if the negative states have become much more vivid, neglecting them may be hard and may cause psychological costs if a decision maker wants to suppress the available information (and avoid evaluations induced by frame (S)). In some cases, when states are particularly vivid, it might even be impossible to abandon frame (S).

A version of our model with switching costs provides a framework of how to think of exogenous events, such as those described above, influencing decision making. Suppose that focus types are initially distributed over the population according to probabilities  $p_E$  (i.e., the share of people who use frame (E)) and  $p_S$  (the share of people who use frame (S)) such that  $p_E + p_S = 1$ . In order to switch one's own type, switching costs  $c_{y,x} \geq 0$  have to be incurred, where  $y \in \{E, S\}$  denotes the prior type and  $x \in \{E, S\}$  the final type. If focus type (S) requires detailed information on the potential states, there may be positive switching costs  $c_{E,S} > 0$  due to information acquisition costs. Switching from (S) to (E) may require the ignoring of possessed information, which may cause some psychological costs,  $c_{S,E} > 0$ , for instance, to "blank" information on the disastrous consequences of an earthquake one has witnessed. Depending on circumstances, those costs could potentially be very large or even infinite. Given these switching costs, a decision maker who initially focuses according to  $y \in \{E, S\}$  solves the following maximization problem

$$\max_{x \in \{S, E\}} \max_{C \in \mathcal{C}} \{U_t^x(c) - c_{y,x}\}.$$

Without or with negligible switching costs this version of the model coincides with our baseline model, and all our results from Section 3 would remain valid. Significant switching costs, however, limit the decision maker's ability to switch the frames and, as a consequence, her overoptimism.

Consider the effects of a recent earthquake on the decision as to whether or not to buy insurance against damages caused by earthquakes. In our setting with switching costs the effects of witnessing the earthquake could be thought of as an instantaneous increase in the share of people,  $p_S$ , who are pre-assigned to frame (S). Alternatively, it could be thought of as increasing the switching costs  $c_{S,E}$  to switch away from focus (S). In either case, the degree of overoptimistic actions (that is, not buying the insurance) is more limited and, hence, the demand for earthquake insurance would increase. In other words, prior to witnessing the earthquake, decision makers rather evaluate according to (E), make an optimistic decision and do not buy the insurance, while after the earthquake, decision makers are more likely to turn to frame (S) and buy insurance. Hence, our model with switching costs can provide a rationale for the observation that insurance demand increases immediately after an earthquake (Lindell and Perry, 2000; Palm, 1995).<sup>19</sup>

Similarly, the model can show why marketing activities by insurance companies should highlight the downside state. In particular, the use of salespeople who can vividly envision the downside states might be an effective tool for insurance companies to increase demand. Initially, in a situation where she is not directly confronted with the potential downside state, and hence switching costs are low, a decision maker may choose an optimistic view by opting for frame (E). In turn, she would ignore the potential adverse states and abstain from an insurance purchase. However, an insurance salesman may succeed in making the downside state vivid which, in our model, could force her into frame (S), or alternatively, contact with the insurance agent may limit overoptimism by increasing the switching cost  $c_{S,E}$ . Then, she may be unable to ignore the adverse states: using frame (E) might only go along with high psychological costs of avoiding thoughts on the potential downside risk. A decision maker may then be locked with frame (S) such that she overweights the future loss states and therefore buys (even excessive) insurance. This distinct prediction on insurance purchases is empirically testable: people not in touch with insurance sellers are hypothesized to be underinsured, whereas people tend to overinsure if they are directly in an insurance sales conversation. Hence, our model may provide a new rationale for the excessive use of sales agents, playing with the fears of decision makers, in the insurance industry.

Furthermore, Bauer *et al.* (2013) show that after the Fukushima nuclear accident prices in the German housing market dropped for those properties which were in short distance to a nuclear power plant. While the authors argue that economic reasons are causal for this drop (closure of nuclear power plants leads to a loss of jobs), our model yields a novel,

---

<sup>19</sup>As a similar example, Dessaint and Matray (2014) find that corporate managers (temporarily) increase cash holdings following a nearby hurricane. This finding is consistent with our model where the occurrence of the hurricane renders the risk from hurricanes salient so that managers tend toward frame (S). As a result, cash holdings are increased as a precautionary measure.



attention-based explanation. Fukushima rendered the adverse events of nuclear disasters particularly salient such that people switched the frame in which they evaluate houses in the catchment area around nuclear power plants. Even though Fukushima did not yield any new information for the likelihood of nuclear disasters in Germany, it induced people to focus on states and therefore to overestimate the risk with which such disasters are realized. The subsequent actions by the German government of switching off some nuclear power plants in the short run and a nuclear phase-out in the long run even reduced the likelihood of nuclear disasters in Germany. However, this reduction in objective probability could not compensate for the increased subjective probability due to people's shift in attention.

The effects of anti-smoking campaigns can be analyzed in a similar fashion. Providing images of blackened smoker lungs, as is planned in the EU, may make decision makers unable to ignore the adverse outcomes which could result from smoking. Thus, we can argue why and how vivid information campaigns work and why governmental figurative anti-smoking campaigns can be particularly effective, which has been confirmed in a number of empirical studies (e.g., Bansal-Travers *et al.*, 2011; Hassan *et al.*, 2008; Munafò *et al.*, 2011). The choice of how to focus can be influenced by vivid information as this may hinder agents from neglecting unpleasant outcomes. Therefore, intervention through information campaigns may affect a decision maker's initial frame (or increase the switching cost) and thereby influence her consumption decision toward less smoking.

## 5 Extensions and limitations

This section extends the results from the base model in the preceding section. We consider a version of our model in which states are correlated between periods and show that in such a setting the main result of overoptimism is robust. Furthermore, we present decision situations where our model does not predict overoptimism. Finally, we argue why our model can also account for several further issues as analyzed by the models which our paper builds on.

### 5.1 Interdependent states

We previously assumed that state realizations at all points in time are independent. This assumption may be too strict for many practical problems. For instance, falling sick with diabetes because of an unhealthy diet at a certain point in time is strongly correlated with suffering from diabetes at a consecutive point in time. Therefore, we generalize our model in order to account for intertemporal correlations. We take the time-horizon  $\mathcal{T}$ , set  $I$  and the random variables  $C_\tau^i$  with the corresponding and the measurable spaces  $(S_\tau, \mathcal{F})$  as introduced in Section 2. Define  $S := \bigoplus_{\tau \in \mathcal{T}} S_\tau$  with the canonical  $\sigma$ -algebra  $\mathcal{F} := \bigotimes_{\tau \in \mathcal{T}} \mathcal{F}_\tau$ .

Consider the probability space  $(S, \mathcal{F}, p)$  for a probability measure  $p$ . Then, we define

$$p(s_\tau) := p(\{(\tilde{s}_t, \dots, \tilde{s}_T) \in S | \tilde{s}_\tau = s_\tau\}) = \sum_{\{(\tilde{s}_t, \dots, \tilde{s}_T) \in S | \tilde{s}_\tau = s_\tau\}} p((\tilde{s}_t, \dots, \tilde{s}_T)).$$

While this gives the formal generalization of our model toward intertemporal interdependence, the examples which we will consider in the following require much less notation. As analyzing general correlations is untractable, we consider a very strong form of correlation and illustrate that our findings are robust with respect to such interdependent state realizations and do not rely on our independence assumptions. Instead, overoptimistic actions are also prevalent in settings where outcomes in different periods are interdependent.

**Harmful consumption and incurable diseases.** Especially for severe diseases like lung cancer or diabetes, the chance to be cured may be tiny. Therefore, we extend our analysis on harmful consumption (Section 3) by assuming that, given a disease (the downside outcome, denoted  $c_{\tau,2}$ ) has occurred in a previous period, then it will last for all future periods. Formally, we assume that  $C_\tau^u = c_{\tau,2}$  induces  $C_{\tau+1}^u = c_{\tau,2}$  with probability one for all  $\tau \in \mathcal{T} \setminus \{T\}$ . If, however, the adverse state has not been realized in the past, then there is a fixed probability  $p$  of becoming ill at each point in time. Therefore, the probability of being ill in period  $\tau \geq t + 1$ , is

$$p^\tau := \sum_{i=t+1}^{\tau} p \cdot (1-p)^{i-t-1}.$$

In particular,  $p^{\tau+1} > p^\tau$  for all  $\tau \in \{t+1, \dots, T-1\}$ . We define  $f_\tau = p^\tau L$  and assume that  $p^{t+n}$  is relatively small for all  $n$ . Then, as before, we obtain  $U_t^S(c^u) < U_t(c^u) < U_t^E(c^u)$ , and a decision maker yields higher focus-weighted utility by choosing frame (E). Hence, our main insight from Section 3 does not change, and the decision maker will choose the risky option as long as

$$Fg(F) > \sum_{\tau=t+1}^T g(f_\tau) f_\tau.$$

This holds in particular if

$$\frac{g(F)}{g(f_T)} > \frac{(T-t)f_T}{F}.$$

Therefore, excessive harmful consumption may arise due to focusing according to frame (E) and does not rely on the state independence assumption which we imposed in the previous sections.

**Persistent business success.** Analogously we can show that overoptimistic actions are also robust with interdependent states in the case of upside risk. For instance, if a newly founded business becomes very successful, then it could be unlikely that it will go bankrupt

within the next periods, but will rather go on to be successful. Consequently, similar to the previous paragraph, we assume that if the favorable outcome (the business success) has occurred in a previous period, then it will last for all future periods. Else, the upside outcome is realized with the fixed probability  $q$ . Analogously to the previous paragraph, the probability with which the favorable outcome occurs in period  $\tau \geq t + 1$  is

$$q^\tau := \sum_{i=t+1}^{\tau} q \cdot (1 - q)^{i-t-1}.$$

Using this definition, one can show, as in Section 3, that by focusing on (S) a decision maker yields a higher focus-weighted utility. That is,  $U_t^S(c^u) > U_t(c^u) > U_t^E(c^u)$ . With focus type (S) the decision maker then chooses to go for the risky option if

$$-H + G \cdot \sum_{\tau=t+1}^T \frac{q^\tau g(G)}{q^\tau g(G) + (1 - q^\tau)g(0)} > 0.$$

A sufficient condition for this to hold is equation (5). Consequently, according to our model, entrepreneurial overoptimism may also occur if business success is persistent over time.

## 5.2 Limitations of overoptimism

While people behave overoptimistically in certain regards, optimistic behavior is not universal. There are natural limitations, which are also captured by our model as we discuss in the following. We examine situations in which distortions of risk are not expected to influence decision making and show why our model is in line with these predictions. In particular, this analysis yields important differences between our model and the model of overoptimism by Mayraz (2011) as the latter cannot account for the limitations of overoptimism we delineate here (see Section 6).

We discuss the choice between a safe option  $c^c$  and a risky option  $c^u$  which represents a *mean preserving symmetric spread* of the safe option. Suppose that the two options yield the same payoffs in all but one period. In that specific period the safe option offers a safe payoff whereas the risky option offers a gamble with a 50% chance of exceeding the payoff of the safe option and a 50% chance of falling short by the same amount. As this represents a mean preserving symmetric spread, both options yield the same expected payoffs.

We will show that in this class of decision situations our model's predictions entirely match the predictions made by expected utility theory, provided that a decision maker's preferences exhibit a diminishing sensitivity with respect to gains and losses.

Let  $c_\tau^u = c_\tau^c$ , for all  $\tau, \tau' \in \mathcal{T} \setminus \{\bar{\tau}\}$ . At  $\tau = \bar{\tau}$ , option  $c^u$  pays  $c_{\bar{\tau}}^c - \alpha$  or  $c_{\bar{\tau}}^c + \alpha$  for some  $\alpha > 0$ , each with a 50% probability. We restrict our analysis to case (1) in which  $[c_{\bar{\tau}}^c - \alpha, c_{\bar{\tau}}^c + \alpha] \subseteq [0, \infty)$  or to case (2) where  $[c_{\bar{\tau}}^c - \alpha, c_{\bar{\tau}}^c + \alpha] \subseteq (-\infty, 0]$ , i.e.  $c_{\bar{\tau}}^c \neq 0$  and  $\alpha$  is sufficiently small. That is, we exclude lotteries which could yield both positive and

negative outcomes with a positive probability. Instead, we investigate lotteries yielding positive outcomes and those yielding negative outcomes separately.<sup>20</sup>

We impose the following assumptions on the agent's utility function. As before, an outcome of zero yields zero utility, and we assume that the utility is strictly monotonic increasing in the outcome. Second, we distinguish whether the agent has a linear utility function and is therefore risk-neutral, or if her utility function satisfies diminishing sensitivity with respect to zero. The latter fulfills our assumptions made in Section 2.1 and implies that she is risk-averse among positive outcomes and risk-seeking among losses.

**Expected utility theory.** Expected utility theory predicts that the risk-neutral agent is indifferent between the two options, whereas an agent obeying diminishing sensitivity is predicted to strictly prefer the safe (risky) option if and only if  $c_{\bar{\tau}}^c > 0$  ( $< 0$ , resp.).

As we will show in the following, our model's predictions are independent of the chosen focus type, and her decisions are in line with expected utility maximization. Therefore, as revealed in the present analysis, overoptimism in our model has plausible limitations. Note that in order to compare utilities derived from the alternatives, we can restrict our analysis to the options' payoffs at the specific point in time  $\bar{\tau}$  since they yield equal payoffs at all other points in time.

**Frame (S).** First, we assume that the decision maker uses representation (S). Given that the agent has a linear utility function, the focus weights on both feasible states are identical and equal  $g(u(c_{\bar{\tau}}^c) - u(c_{\bar{\tau}}^c - \alpha)) = g(u(c_{\bar{\tau}}^c + \alpha) - u(c_{\bar{\tau}}^c))$ . Consequently, both options' focus-weighted utilities match the actual consumption utilities and the agent is indifferent between both alternatives, i.e.,  $U(c^c) = U^S(c^c) = U^S(c^u)$ .

Next, we consider a utility function which satisfies diminishing sensitivity with respect to zero. In case (1), the focus weight on  $c^u$ 's relative downside is particularly large,  $g(u(c_{\bar{\tau}}^c) - u(c_{\bar{\tau}}^c - \alpha)) > g(u(c_{\bar{\tau}}^c + \alpha) - u(c_{\bar{\tau}}^c))$ , so that  $U^S(c^u) < U^S(c^c)$ . This is reversed in case (2): the focus weight on the relative upside becomes particularly large, which triggers preferences in favor of  $c^u$ , i.e.,  $U^S(c^u) > U^S(c^c)$ . Therefore, given a focus on states, the risky option is preferred if the agent has to choose between losses while the safe alternative is preferred if she chooses between safe and risky gains. Depending on the underlying utility function, these predictions exactly meet expected utility maximization. Note that the agent, however, overestimates in each case the difference in utility between the two available options. If the risky option yields the higher focus-weighted utility, then its relative upside is overvalued, whereas if the safe option is preferred, then the risky option's relative downside is overvalued.

---

<sup>20</sup>In the following we detect two different effects in the positive and in the negative domain. If  $sign(c_{\bar{\tau}}^c + \alpha) \neq sign(c_{\bar{\tau}}^c - \alpha)$ , then both then standard theory's and our model's predictions crucially depend on which of the two effects is dominant. Dominance depends on the curvature of the utility function in the respective domain, which we are agnostic about here. Thus, we restrict our analysis to those cases where predictions by standard theory and by our model are unambiguous.

**Frame (E).** Second, we analyze utilities if the agent incorporates frame (E). With a linear utility function both alternatives yield the same expected utilities at all points in time. Therefore, the agent is indifferent between both options. If the utility function satisfies diminishing sensitivity, the focus-weighted utilities are as follows. In case (1), the safe option  $c^c$  gives a higher expected utility at  $\bar{\tau}$  due to risk-aversion. Provided  $T > t$ , both options' focus weights are larger at  $\bar{\tau}$  than at all other points, so that the safe option is preferred while the difference in utility between the two options is overrated. The reverse holds in case (2) where the risky option is preferred and the difference in utility between the two options is overrated as  $U^E(c^u) - U^E(c^c) > U(c^u) - U(c^c)$ . Therefore, our model predicts rational choice if the agent focuses on expected utilities.

**Proposition 5** *Suppose a decision maker chooses between a safe option and a riskier alternative, which is a mean preserving symmetric spread of the safe option. Then, the following predictions arise both from expected utility maximization and from our model. A risk-neutral decision maker is indifferent between both options. A risk-averse decision maker strictly prefers the safe option over the risky alternative, while a risk-seeking decision maker strictly prefers the risky option.*

**Remark.** Note that the preceding analysis carries over to slightly more general cases. First, it holds true if the same symmetric risk was added at more than one point in time (given that the risky option yields at all these points the same expected payoff). Second, it holds true if we add a little symmetric risk to the safe option in period  $\bar{\tau}$ . In this case, the risky option is riskier as the symmetric spread in period  $\bar{\tau}$  is larger for the risky than for the safe option. Then, our model meets predictions by rational choice as a risk averse decision maker will opt for the safer and a risk-seeking decision maker will opt for the riskier option.

To sum up, we have shown that our model does not predict deviations from expected utility maximization in such cases where the latter is most likely to hold true. Therefore, we analyzed a setting in which symmetric risk was added to a safe option as this is one of the most robust settings where expected utility theory is typically assumed to be valid since such a choice problem exactly reflects the definition of risk aversion.<sup>21</sup> Here, our model does not predict overoptimistic actions. Instead, an agent following our model opts for the safe option if and only if it is preferred by a rational agent.

---

<sup>21</sup>Risk aversion can be 1-1 identified with preferences for second-order stochastic dominating options, see for example Hadar and Russell (1969). Therefore, in the simplest test for risk-aversion, subjects choose between (1) a monetary lottery with symmetric risk and (2) its expected payoff. Here, (1) is the simplest lottery which is second-order stochastically dominated by (2). For instance, Kahneman (2011) lists this example in order to define risk aversion. Therefore, all studies which test for risk aversion insinuate that the revealed preference approach is valid in such simple setups and that information about one's true utility function can be elicited through this procedure. Our model also predicts that decision makers reveal their true preferences in such setups.

### 5.3 Further applications

Our model also reproduces other major insights from Köszegi and Szeidl (2013) in the domain of riskless decision making. For instance, according to frame (E) people are more likely to make short-term than long-term commitments with respect to abstaining from harmful consumption, saving for retirement or with respect to exercising. That is because short-term commitments yield only minor future improvements which do not attract much attention, while long-term commitments may result in substantial concentrated improvements which tend to be overrated according to the bias toward concentration. As a consequence, decisions are present-biased with respect to lifestyle choices, but not with respect to concentrated choices between an immediate outcome and a future outcome. Likewise, future-biased decisions as predicted in Köszegi and Szeidl (2013), for instance, if people overcommit to concentrated goals, are in line with our model. This means that people may *ex ante* commit to exercising daily in order to achieve the large goal of completing a marathon, which may be excessive with respect to the expected payoff. As in such riskless setups there are no states to focus on, frame (S) induces rational valuations. Frame (S) will not be selected in equilibrium in the preceding examples as in contrast to frame (E) it does not allow subjects to focus on (and, therefore, to overweight) an especially favorable happening. Thus, our model also reproduces those choice patterns which stand in interesting contrast to models of present-biased decision making.

In addition, the attribute space does not have to consist of time periods in which outcomes can be achieved. Instead, attributes could represent characteristics of available goods such as *quality* or *price*. In such a setup, our model also generates hypotheses raised by theories of distorted attention with respect to riskless decision making. For instance, papers on mental accounting argue that a consumer's willingness to pay for a certain object depends on the choice context (Bordalo *et al.*, 2013; Thaler, 1980; Thaler, 1999). In particular, Bordalo *et al.* (2012) argue that uniform upward price shifts of vertically differentiated goods may raise the demand for the pricier item with the higher quality. They provide the example of a consumer choosing between a cheap Australian wine and a more expensive French wine of a higher perceived quality. While she may buy the cheap wine at a wine store, she may switch to the more expensive wine at a classy restaurant (even if her consumption utility is not altered by the different consumption location) if both wines' prices are, compared to the store, uniformly upward shifted. As prices are less salient in the latter scenario than in the former, consumers tend to focus on prices in the former, but on quality in the latter scenario. The crucial driver for this choice pattern is diminishing sensitivity of salience with respect to an endogenous reference point which equals the average value of the respective attribute (quality or price) among all options in the choice set. We, instead, incorporate diminishing sensitivity of utility with respect to the fixed reference point of zero, which yields the same predictions as Bordalo *et al.* (2013) for all their major applications. For instance, our frame (E) also predicts less attention for the price dimension after a uniform upward price shift as the difference in utilities

gained from the attribute *price* becomes smaller with a rising price level. Therefore, in our framework the substitution toward the pricier item can be rationalized as follows. Since no risk is involved, frame (S) gives the rational valuation of the decision situation while frame (E) shifts attention either toward the price or the quality. If all prices are low, their difference is particularly salient and also frame (E) does not allow their partial neglect. Therefore, the consumer maximizes her expected utility by choosing the Australian wine. If, however, prices are uniformly and sufficiently increased, frame (E) guides attention toward quality, such that the consumer overweights the French wine's quality advantage, opts for it and thereby maximizes her focus-weighted utility. Consequently, our model offers an explanation for such an inconsistent choice pattern which violates rational choice theory. Therefore, our model also unifies the effects of attention distortion beyond the domains of decision making under risk or under intertemporal trade-offs.

## **6 Related literature on self-servingness, overconfidence, and framing**

Our model shares the predictions of other models, in which decision makers cannot only choose a good, but in which they can also decide on their expectations, their beliefs or their framing of a given decision problem. In this section, we will briefly point out differences between our model and these related models.

Mayraz (2011)'s model of wishful thinking assumes that a decision maker's beliefs depend crucially on her interests and how her interests bias the processing of information. A single parameter measures if a decision maker is optimistic or pessimistic. In the case of optimism, the decision maker maximizes a distorted utility function, according to which probability weights are upward distorted for those events which are favorable for the decision maker, while they are downward distorted for unfavorable events. This, however, predicts that overoptimism is, for optimistic individuals, universal. Therefore, this model cannot account for rational, risk-averse behavior as predicted by our model (see Section 5.2).

According to Brunnermeier and Parker (2005), decision makers experience anticipatory utility, such that decision makers may benefit from overoptimistic beliefs as gains in anticipatory utility may outweigh losses resulting from overoptimistic actions. Both Brunnermeier and Parker and our model predict rational decision making only in settings where decision makers choose between a safe option and a mean preserving symmetric spread. In general, Brunnermeier and Parker predict a preference for risky options with positively skewed distributions as these allow for overoptimistic beliefs concerning highly positive, but unlikely events. Thereby, their model can account for gambling, lottery purchases and related phenomena, which have also been rationalized by our mechanism. A positive skew induces decision makers in our model to choose frame (S) and to thereby overweight unlikely, salient outcomes as these are favorable if the distribution is posi-

tive. In contrast to our model, however, a negative skew induces decision makers to be pessimistic in the model by Brunnermeier and Parker. We predict that decision makers are induced to use frame (E) for decision situations in which the risky option is negatively skewed, so that decisions are either overoptimistic (Section 3) or rational (Section 7). Consequently, the model by Brunnermeier and Parker cannot account for overoptimism with respect to harmful consumption. Even further, our model can describe why external forces, which yield no valuable information for the own decision situation, can induce overcautious decision making. Neither Mayraz (2011) nor Brunnermeier and Parker (2005) can rationalize such drastic changes in the processing of risk which we have derived in Section 4.

Empirical evidence for the presence of optimal expectations and the model by Brunnermeier and Parker, such as by Oster *et al.* (2013), is also in line with our (extended) model as the findings by Oster *et al.* could be interpreted as follows. Individuals who are untested for the hereditary Huntington’s disease engage in overoptimistic actions as they neglect the risk of being ill. In contrast, if individuals receive the information that they are ill, the individuals’ attention is guided toward states as implied by frame (S), which triggers very different decision making. Therefore, our central assumption of self-serving framing is also supported by the empirical literature on manipulable expectations.

Moreover, models of risk-seekingness cannot produce our results. First, such models cannot account for the plausible limitations of overoptimism (Section 5.2). Second, such models cannot explain a shift in risk attitudes through external events as analyzed in Section 4.

## 7 Attention in simple lotteries

In this section, we investigate central, long-known phenomena of binary choice under risk (see, for instance, Kahneman and Tversky, 1979), which, among others, salience theory of choice under risk as outlined in Bordalo *et al.* (2012) can account for. We will investigate the Common-Ratio Allais Paradox (CRAP), the Common-Consequence Allais Paradox (CCAP) and additionally Risk Attitudes under Salience (RAS). As these decisions involve only a single future period, these phenomena do not relate to intertemporal decision making. We apply the simplified one-stage version of our model to these decision situations and contrast its predictions with those of salience theory as presented in Bordalo *et al.* (2012). In line with Bordalo *et al.* (2012), in this section we denote the choice options as *lotteries* which are uniquely described by the outcomes they yield in all states. We assume that the number of states in the state space is  $n \in \mathbb{N}$ . An agent is assumed to face the binary choice between the two lotteries  $L_1$  and  $L_2$ , where  $L_i := (x_i^1, p_1; \dots; x_i^n, p_n)$  is defined as the lottery which realizes outcome  $x_i^k$  with probability  $p_k$  for all  $1 \leq k \leq n$  and  $i = 1, 2$ .

For each of the choice situations under consideration, we derive predictions by our model and by Bordalo *et al.* (2012). In order to compare the predictions, we impose a



specification, described in subsection 7.1, such that in both approaches the same states are rendered salient. Our choice tasks comprise some questions where predictions by our model and by Bordalo *et al.* (2012) are perfectly in line and some where the predictions (mainly for lotteries in the loss domain) are divergent.

In order to test both models' predictions against each other, we conducted two surveys ( $N = 173$ ). We tested for two versions of each of these three choice situations, one involving gains ( $N = 90$ ) and the other one involving losses ( $N = 83$ ). In each task, subjects had to choose from a choice set  $C$  containing lotteries which all yield either positive or negative outcomes. For this investigation, we define  $-L$  to be lottery  $L$  with the payoffs' signs reversed, i.e., all payoffs are multiplied by  $-1$ . Similarly,  $-C$  is defined to consist of all lotteries in  $C$  where all lotteries' outcomes have inverted signs. To our knowledge, these paradoxes with negative signs have not been tested before.

## 7.1 Our model's relation to salience theory of choice under risk

In this paragraph we will argue why for binary choices which yield payoffs in a single period the focus according to (S) is largely equivalent to the susceptibility to salience in the sense of Bordalo *et al.* (2012). First, note that frame (E) yields the rational valuation of an option (and is therefore equivalent to frame (R)). As there are no different points in time which may attract different degrees of attention, agents focusing according to (E) necessarily evaluate their choices rationally. Second, however, focusing according to (S) yields distorted valuations of risky options as it represents a bias of risk perception. While Bordalo *et al.* (2012) measure the salience of a lottery's outcome via a *salience function*  $\sigma$ , in our setup weights are provided by the focusing function  $g$ . We will show that the underlying assumptions according to which the relative salience of an option's outcome is assessed are the same for the two approaches. Therefore, if an outcome can be presumed to be salient according to Bordalo *et al.* (2012), then it can also be presumed to be salient according to our approach (and vice versa). This means that both approaches produce the same ranking of states' relative salience.

Moreover, the functional forms yielding the focus weights are not significantly different between the distortion mechanisms of the two approaches. To determine focus weights, Bordalo *et al.* (2012) do not use the absolute values of the outcomes' salience, but only the salience ranking of outcomes. Focus weights are assigned to outcomes according to a discrete function of a salience parameter  $\delta \in [0, 1]$ , which indicates the susceptibility for salience of a decision maker (the lower  $\delta$ , the higher the decision maker's susceptibility). However, Bordalo *et al.* (2012) also provide a continuous approach in their section III.B, which does not alter their qualitative results, but according to which the explicit degree of an outcome determines to which degree its valuation is distorted. This mechanism is closely related to our approach according to which focus weights are generated through the continuous, strictly monotonic increasing focusing function  $g$ . How focus weights result from the explicit degree of salience is then equivalent in both approaches

as in each case the weights equal the explicit degree of salience, normalized via the probability weighted sum of all focus weights. In particular, in both models the focus weight on a lottery's outcome is proportional to its salience. Therefore, given the equivalence of relative salience, our model with focus (S) is able to explain the same phenomena as the model of Bordalo *et al.* (2012).

We proceed by deriving the two approaches' equivalence of the salience rankings of states in more detail. All key properties determining the salience of outcomes according to Bordalo *et al.* (2012), as listed in the following, are one to one mirrored by our model.

First, Bordalo *et al.* (2012) incorporate the linear utility function such that  $L_i$ 's payoff  $x$  gives an agent utility  $x$ . Second, to assess how salient the outcomes of a lottery in certain states are, the authors define a salience function  $\sigma$  through the properties of ordering, diminishing sensitivity and reflection as introduced below. In contrast, our model comprises a utility function  $u(\cdot)$  satisfying diminishing sensitivity with respect to zero and a strictly monotonic increasing focusing function determining a state's salience level. Both approaches give rise to the same salience ranking of a lottery  $L_i$ 's outcomes. To see that, we compare both models' functions assigning a pair of outcomes a degree of salience, i.e.,  $\sigma : \mathbb{R}^2 \rightarrow \mathbb{R}$  and  $g \circ u : \mathbb{R}^2 \rightarrow \mathbb{R}$  with  $g \circ u(x, y) := g(|u(x) - u(y)|)$ . According to Bordalo *et al.* (2012),  $L_i$ 's outcome  $x_s^i$  is more salient than  $x_{\tilde{s}}^i$  if and only if  $\sigma(x_s^i, x_s^{-i}) > \sigma(x_{\tilde{s}}^i, x_{\tilde{s}}^{-i})$ . Our model assigns  $x_s^i$  a higher degree of salience than  $x_{\tilde{s}}^i$  if and only if  $g \circ u(x_s^i, x_s^{-i}) = g(|u(x_s^i) - u(x_s^{-i})|) > g(|u(x_{\tilde{s}}^i) - u(x_{\tilde{s}}^{-i})|) = g \circ u(x_{\tilde{s}}^i, x_{\tilde{s}}^{-i})$ . Therefore, it remains to show that  $\sigma$  and  $g \circ u$  share the same basic properties.

Bordalo *et al.* (2012) define the salience of state  $s$  for lottery  $L_i$ , with  $i = 1, 2$ , as a continuous function  $\sigma(x_s^i, x_s^{-i})$  that satisfies three conditions:

- 1) *Ordering.* If for states  $s, \tilde{s} \in S$  the interval  $[\min\{x_s^i, x_s^{-i}\}, \max\{x_s^i, x_s^{-i}\}]$  is a strict subset of  $[\min\{x_{\tilde{s}}^i, x_{\tilde{s}}^{-i}\}, \max\{x_{\tilde{s}}^i, x_{\tilde{s}}^{-i}\}]$ , then  $\sigma(x_s^i, x_s^{-i}) < \sigma(x_{\tilde{s}}^i, x_{\tilde{s}}^{-i})$ .
- 2) *Diminishing sensitivity of salience.* Given  $x_s^j > 0$  for  $j = 1, 2$  and given any  $\epsilon > 0$ , then  $\sigma(x_s^i + \epsilon, x_s^{-i} + \epsilon) < \sigma(x_s^i, x_s^{-i})$ .
- 3) *Reflection.* For all  $s, \tilde{s} \in S$  such that  $x_s^j, x_{\tilde{s}}^j > 0$  for  $j = 1, 2$ , we have  $\sigma(x_s^i, x_s^{-i}) < \sigma(x_{\tilde{s}}^i, x_{\tilde{s}}^{-i})$  if and only if  $\sigma(-x_s^i, -x_s^{-i}) < \sigma(-x_{\tilde{s}}^i, -x_{\tilde{s}}^{-i})$ .

The ordering property means that lotteries  $L_i$  and  $L_{-i}$ 's outcomes in a certain state increase in salience if their range of outcomes is increased. This property is exactly mirrored by our assumption that the focusing function and the utility function are both strictly monotonic increasing. In our model, therefore, the following equivalent condition holds.

- 1) *Monotonicity.* If for states  $s, \tilde{s} \in S$  the interval  $[\min\{x_s^i, x_s^{-i}\}, \max\{x_s^i, x_s^{-i}\}]$  is a strict subset of  $[\min\{x_{\tilde{s}}^i, x_{\tilde{s}}^{-i}\}, \max\{x_{\tilde{s}}^i, x_{\tilde{s}}^{-i}\}]$ , then  $g \circ u(x_s^i, x_s^{-i}) < g \circ u(x_{\tilde{s}}^i, x_{\tilde{s}}^{-i})$ .

The second property, diminishing sensitivity, states that a uniform increase in all positive payoffs makes a state less salient. In our setup, this property is incorporated through the utility function  $u$  such that the following holds.

- 2) *Diminishing sensitivity of utility.* Given  $x_s^j > 0$  for  $j = 1, 2$  and given any  $\epsilon > 0$ , then  $u(x_s^i + \epsilon, x_s^{-i} + \epsilon) < u(x_s^i, x_s^{-i})$  and  $g \circ u(x_s^i + \epsilon, x_s^{-i} + \epsilon) < g \circ u(x_s^i, x_s^{-i})$ .<sup>22</sup>

The reflection property denotes the fact that switching all outcomes' signs does not render other states salient such that the salience ranking of the different states and outcomes is preserved. While this property is not a necessary implication of our modeling, it can be incorporated, for example, by the assumption of a point-symmetric utility function. In that case,  $|u(x_s^i) - u(x_s^{-i})| = |u(-x_s^i) - u(-x_s^{-i})|$  holds. Therefore, such a utility function yields

- 3) *Reflection and point-symmetric utility.* For all  $s, \tilde{s} \in S$  such that  $x_s^j, x_{\tilde{s}}^j > 0$  for  $j = 1, 2$  we have  $g \circ u(x_s^i, x_s^{-i}) < g \circ u(x_{\tilde{s}}^i, x_{\tilde{s}}^{-i})$  if and only if  $g \circ u(-x_s^i, -x_s^{-i}) < g \circ u(-x_{\tilde{s}}^i, -x_{\tilde{s}}^{-i})$ .

Thus, we impose the following specification for our subsequent analysis:

**Specification.** *With focus (S), suppose that an agent has to choose between two risky options  $L_1$  and  $L_2$ . If an option  $L_i$ 's outcome  $x_s^i$  is relatively salient in the sense of Bordalo et al. (2012) and therefore overweighted compared to outcome  $x_{\tilde{s}}^i$  for states  $s, \tilde{s}$ , then this outcome is relatively salient and overweighted according to our model with focus type (S). Furthermore, an agent values  $L_i$  higher than  $L_{-i}$  according to Bordalo et al. (2012) if and only if she values  $L_i$  higher than  $L_{-i}$  according to our model.*

## 7.2 Common ratio Allais paradox

One well-known paradox of choice under risk, which prospect theory and salience theory can account for, is the common ratio Allais paradox. It indicates a shift toward risk seeking as the probability of winning falls, which violates expected utility theory. Bordalo et al. (2012) test for this paradox by investigating choices between  $L_1(\bar{\pi}) = (6000, \bar{\pi})$  and  $L_2(\pi) = (3000, \pi)$  for different values of  $\pi$  and  $\bar{\pi}$  where absolute numbers indicate US-\$ in their setup and € in our setup. Agents have to choose from  $C_A = \{L_1(0.001), L_2(0.002)\}$  and from  $C_B = \{L_1(0.45), L_2(0.9)\}$  one option each. If an agent is susceptible to salience to a sufficient degree in the sense of Bordalo et al. (2012), then she will choose  $L_1(0.001)$  and  $L_2(0.45)$  as the authors argue. This choice pattern is known as the common ratio Allais paradox. In contrast, expected utility theory predicts consistent choices of either  $L_1(0.001)$  and  $L_1(0.45)$  or  $L_2(0.002)$  and  $L_2(0.9)$ .

In both decision situations, four states could be realized, which are, ranked from the most to the least salient,  $s_1 = (6000, 0)$ ,  $s_2 = (0, 3000)$ ,  $s_3 = (6000, 3000)$  and  $s_4 = (0, 0)$ . For each tuple, the first number denotes  $L_1$ 's outcome and the second number denotes  $L_2$ 's outcome. This salience ranking is unique both for salience theory of choice under risk and for our model (provided a strictly monotonic increasing focusing function and a utility function satisfying diminishing sensitivity).

<sup>22</sup>The latter follows from our assumption that  $g$  is strictly monotonic increasing.

First, we analyze choices from  $C_A$ . Under plausible and weak assumptions (for details, see Appendix B) the ordering of utilities is as follows

$$U^S(L_1(0.001)) > U^S(L_2(0.002)) > U^E(L_2(0.002)) > U^E(L_1(0.001)).$$

The first comparison  $U^S(L_1(0.001)) > U^S(L_2(0.002))$  holds, as Bordalo *et al.* (2012) show, if an agent's susceptibility to salience is sufficiently strong since  $(6000, 0)$  is very salient and since distortions of probabilities are especially large for unlikely states. In order to see why the second inequality  $U^S(L_2(0.002)) > U^E(L_2(0.002))$  holds, remember that a focus on the expected utilities (E) is equivalent to the rational assessment as there is only one point in time.  $L_2$ 's upside of winning 3000 is relatively salient, which results in an overall overrating of  $L_2$ . Therefore, the focus-weighted utility which  $L_2$  yields is higher under a focus on states (S) than under (E). Finally,  $U^E(L_2(0.002)) > U^E(L_1(0.001))$  holds since the agent is assumed to be risk-averse in the positive domain. This follows directly from the assumption of diminishing sensitivity which her actual utility function is assumed to satisfy. To sum up, Bordalo *et al.* (2012) and our model both predict that a decision maker chooses  $L_1(0.001)$  in order to maximize her focus-weighted utility.

Due to the monotonicity of the utility function, this ranking is reversed if all positive outcomes are substituted by negative outcomes such that

$$U^S(-L_1(0.001)) < U^S(-L_2(0.002)) < U^E(-L_2(0.002)) < U^E(-L_1(0.001)).$$

Consequently, our model predicts that agents choose focus type (E) and choose lottery  $-L_1(0.001)$ , whereas salience theory predicts a choice of  $-L_2(0.002)$ . If overoptimism guides a decision maker's decision making in this choice context, then she abstains from focusing on the (adverse) states and therefore meets predictions by rational choice, but not by Bordalo *et al.* (2012).

Second, we analyze choices from  $C_B$ . In order to predict decisions we only need the two inequalities

$$U^E(L_2(0.9)) > U^E(L_1(0.45)).$$

and

$$U^S(L_2(0.9)) > U^S(L_1(0.45)).$$

The latter inequality,  $U^S(L_2(0.9)) > U^S(L_1(0.45))$ , holds due to preferences for relatively safe outcomes due to risk aversion and is satisfied by the model of Bordalo *et al.* (2012). Furthermore,  $U^E(L_2(0.9)) > U^E(L_1(0.45))$  clearly follows from the agent's risk aversion in the gain domain. Consequently, our model predicts that subjects choose lottery  $L_2(0.9)$ , no matter their focus type. This meets the prediction by salience theory. For inverted signs, both models predict that  $-L_1(0.45)$  is chosen.

**Prediction 1** *Our model and salience theory of choice under risk predict that subjects choose  $L_1(0.001)$  and  $L_2(0.9)$ . For losses, our model predicts choices of  $-L_1(0.001)$  and  $-L_1(0.45)$ ,*

while salience theory predicts that  $-L_2(0.002)$  and  $-L_1(0.45)$  are chosen.

Our model predicts that for positive outcomes agents decide in line with Bordalo *et al.* (2012) since the options are relatively overrated given focus (S). This, however, changes for losses. Here, the options' downsides may be more salient and therefore overweighted, such that the decision maker may decide to focus according to (E), i.e., to choose the rational option. If susceptibility to salience in the sense of Bordalo *et al.* (2012) is accompanied not by optimism but by pessimism, i.e., with an undervaluation of the alternatives, then our model's predictions deviate from the ones in Bordalo *et al.* (2012). Consequently, in our model only, decision makers are predicted to value risky options rationally as long as salience distortions do not result in overvaluations.

**Experimental findings.** For gains, only a minority maximizes expected utility by choosing twice  $L_1$  or twice  $L_2$ . The most frequent choice pattern, which is chosen by two-thirds of the participants, meets the Common ratio Allais paradox and the predictions both by our model and by salience theory of choice under risk. As predicted by us, choice patterns are fundamentally different for losses. Overall, choices are much more heterogeneous than with gains. Our prediction fares slightly better than the one by Bordalo *et al.* (2012). About 30% of the participants choose according to our prediction while a bit more than 25% choose according to Bordalo *et al.* (2012). Both theories can, hence, account for a significant fraction of observed behavior.

	$L_1^{0.45}$	$L_2^{0.9}$	
$L_1^{0.001}$	14.4	66.7	81.1
$L_2^{0.002}$	1.1	17.8	18.9
	15.6	84.4	N = 90

(a) Gains (choices in %)

	$-L_1^{0.45}$	$-L_2^{0.9}$	
$-L_1^{0.001}$	29.6	18.5	48.2
$-L_2^{0.002}$	25.9	25.9	51.9
	55.6	44.4	N = 81

(b) Losses (choices in %)

Table 1: Common ratio Allais paradox

### 7.3 Common consequence Allais paradox

The common consequence Allais paradox is another well-known violation of expected utility theory. One explanation is the so-called certainty effect, which means that preferences are discontinuous over certain and uncertain outcomes (Andreoni and Sprenger, 2010). Equivalently, the probability weighting function can be assumed to be discontinuous at  $p = 1$ . Salience effects can also provide a plausible explanation for the observed behavior. To test for this paradox, Bordalo *et al.* (2012) investigate the following decision situations. An agent has to choose between  $L_1(z) = (2500, 0.33; 0, 0.01; z, 0.66)$  and  $L_2(z) = (2400, 0.34; z, 0.66)$  for different values of  $z$ . In particular, agents have to choose one option from  $C_A = \{L_1(2400), L_2(2400)\}$  and one from  $C_B = \{L_1(0), L_2(0)\}$ . According to Bordalo *et al.* (2012), the options  $L_2(2400)$  and  $L_1(0)$  are chosen since the different values of  $z$  render different states salient. This phenomenon, the common consequence

Allais paradox, represents a violation of expected utility theory since choices are not consistent.

In decision situation *A* three states are feasible, which are, ranked from the most to the least salient state,  $s_1 = (0, 2400)$ ,  $s_2 = (2500, 2400)$  and  $s_3 = (2400, 2400)$ . In situation *B* four states are feasible, which are, ordered by their salience,  $s_1 = (2500, 0)$ ,  $s_2 = (0, 2400)$ ,  $s_3 = (2500, 2400)$  and  $s_4 = (0, 0)$ . Note that both salience rankings are unique both according to salience theory of choice under risk and to our model.

First, we analyze choices from  $C_A$ . The ordering of focus-weighted utilities is as follows:

$$U^E(L_1(2400)) > U^S(L_2(2400)) = U^E(L_2(2400)) > U^S(L_1(2400)).$$

Since  $L_2(2400)$  represents a safe option, risk distortions do not influence the valuation of  $L_2(2400)$  and therefore the equality  $U^S(L_2(2400)) = U^E(L_2(2400))$  clearly holds. If decision makers' susceptibility to salience is sufficiently strong, then the comparison  $U^S(L_2(2400)) > U^S(L_1(2400))$  holds as Bordalo *et al.* (2012) also argue. This results from a strong distortion of the small probability of winning zero which induces choosing  $L_2(2400)$ .<sup>23</sup> The inequality  $U^E(L_1(2400)) > U^E(L_2(2400))$  holds if the agent's risk aversion is not extremely strong.<sup>24</sup> Hence, Bordalo *et al.* (2012) predict that a decision maker will choose  $L_2(2400)$  since she wants to avoid the very salient zero outcome in state  $(0, 2400)$  while we predict that she will choose  $L_1(2400)$ . With negative signs both approaches hypothesize that option  $-L_1(2400)$  is chosen as here focus (S) induces that favorable states are salient and overweighted.

Next, we analyze choices from  $C_B$ . For a local thinker, the ordering of utilities is

$$U^S(L_1(0)) > U^S(L_2(0))$$

and

$$U^E(L_1(0)) > U^E(L_2(0)).$$

Here, the first comparison  $U^S(L_1(0)) > U^S(L_2(0))$  clearly holds as Bordalo *et al.* argue: the gain of 2500 is more salient than the potential gain of 2400 and the probability difference between these potential gains is only very small (0.33 vs. 0.34). Finally, under the condition that the agent is not extremely risk averse (which has also been assumed in decision situation *A*), the comparison  $U^E(L_1(0)) > U^E(L_2(0))$  holds.<sup>25</sup> Therefore, Bordalo *et al.* and our model, no matter the focus type, predict that  $L_1(0)$  is chosen in the case of positive outcomes. If signs are reversed, then both approaches predict that  $-L_2(0)$  is chosen.

<sup>23</sup>Based on a rough calibration, Bordalo *et al.* (2012) incorporate the salience function  $\sigma(x, y) = |x - y|/(|x| + |y| + 0.1)$ . Then, agents reveal the common consequence Allais paradox, i.e., they have  $U^S(L_2(2400)) > U^S(L_1(2400))$  but  $U^S(L_1(0)) > U^S(L_2(0))$  if  $\delta < 0.73$ , which means that their susceptibility to salience has to be sufficiently strong.

<sup>24</sup>Given utility function  $u(x) = x^\alpha$ , solving  $2400^\alpha = 0.33 \cdot 2500^\alpha + 0.66 \cdot 2400^\alpha$  shows that the inequation holds for all  $\alpha > 0.73$ .

<sup>25</sup>Again, provided the utility function  $x^\alpha$ , this inequation holds for all  $\alpha > 0.73$ .

**Prediction 2** Our model predicts that a decision maker will choose  $L_1(2400)$  and  $L_1(0)$ , while salience theory of choice under risk predicts that subjects will choose  $L_2(2400)$  and  $L_1(0)$ . For losses, both models predict choices of  $-L_1(2400)$  and  $-L_2(0)$ .

In contrast to the previous paradox, our model’s predictions deviate from Bordalo *et al.* (2012)’s exclusively for positive outcomes since  $L_1(2400)$ ’s downsides are relatively salient under a focus on (S). Rationally, however, this alternative represents the preferable option. Therefore, our model predicts rational and consistent choices for positive outcomes.

**Experimental findings.** For gains, 45.6% of the observed choice patterns obey the common consequence Allais paradox and can be explained by salience theory of choice under risk, while our model accounts for a share of 40%. This finding is in line with the literature as the common consequence Allais paradox usually represents the modal choice pattern (but not necessarily the choice pattern chosen by a majority of decision makers; for a survey, see Machina (1989)). For losses, a share of 50% is explained by both models. While salience theory of choice under risk performs slightly better than our approach, for positive outcomes the common consequence paradox is not as widespread as the common ratio paradox due to the very high share of rational agents, which is predicted by our model. Experimental support is especially strong for such a phenomenon, i.e., for the common ratio paradox, where predictions by Bordalo *et al.* (2012) and our model agree, which could also be regarded as supportive of our approach.

	$L_1(2400)$	$L_2(2400)$	
$L_1(0)$	40.0	45.6	85.6
$L_2(0)$	2.2	12.2	14.4
	42.2	57.8	N = 90
(a) Gains (choices in %)			
	$-L_1(2400)$	$-L_2(2400)$	
$-L_1(0)$	30.5	2.4	32.9
$-L_2(0)$	50.0	17.1	67.1
	80.5	19.5	N = 82
(b) Losses (choices in %)			

Table 2: Common consequence Allais Paradox

#### 7.4 Risk attitudes under salience

The following phenomenon reveals how salience and risk attitudes are intertwined. Define  $L_1(x) := (1+x, 0.95; 381+x, 0.05)$  and  $L_2(x) := (20+x, 1)$ . Agents have to choose one option from  $C_A = \{L_1(0), L_2(0)\}$  and one alternative from  $C_B = \{L_1(300), L_2(300)\}$ . In Section IV, Bordalo *et al.* (2012) predict that local thinkers will choose  $L_2(0)$  and  $L_1(300)$  (“risk paradox”). Booth and Nolen (2013) yield experimental support for this hypothesis. This risk paradox is related to the common consequence Allais paradox since the options in the sets  $C_A$  and  $C_B$  differ only with respect to a payoff which is added to all options. However, here this added component (an amount of 300) is not stochastic, but

certain. Agents revealing the hypothesized choice pattern seem to have inconsistent risk preferences since a uniform upward shift of payoffs enhances the agent's risk-seekingness. According to Bordalo *et al.* and Booth and Nolen the state  $s_1 = (1, 20)$  is more salient than  $s_2 = (381, 20)$  in decision situation  $A$ , while in  $B$  state  $s_1 = (681, 320)$  is more salient than  $s_2 = (301, 320)$ . Given these rankings, salience theory accounts for choices of  $L_2(0)$  and  $L_1(300)$  for all  $\delta \in (0, 1)$ . However, these salience rankings are, in contrast to the salience rankings in the two previous subsections, not unique. If diminishing sensitivity is not extremely strong, then the salience ranking reverses in decision situation  $A$  (whereas the ranking of states in context  $B$  also holds if diminishing sensitivity is very strong).

First, we analyze choices from  $C_A$ . If the focusing effect is strong enough, the following ordering holds

$$U^S(L_2(0)) = U^E(L_2(0)) > U^E(L_1(0)) > U^S(L_1(0)).$$

Here,  $U^S(L_2(0)) = U^E(L_2(0))$  clearly holds since both focusing types yield the rational valuation for a safe option. The second inequality  $U^E(L_2(0)) > U^E(L_1(0))$  follows from the agent's risk aversion in the positive domain and  $U^E(L_1(0)) > U^S(L_1(0))$  follows from the assumption that  $(1, 20)$  is more salient than  $(381, 20)$ . Consequently, our model and salience theory of choice under risk predict that a local thinker will choose  $L_2(0)$  and  $-L_1(0)$ .

If the salience ranking is reversed in situation  $A$ , i.e., if diminishing sensitivity is not extremely strong, then the ranking of focus weight utilities is

$$U^S(L_2(0)) = U^E(L_2(0)) < U^E(L_1(0)) < U^S(L_1(0)).$$

In that case, both salience theory and our model predict that  $L_1(0)$  and  $-L_2(0)$  will be chosen.

Second, we analyze choices from  $C_B$ , for which the salience ranking is evident. The focus-weighted utilities can be ranked as follows:

$$U^S(L_1(300)) > U^S(L_2(300)) = U^E(L_2(300)) > U^E(L_1(300)).$$

Here,  $U^S(L_1(300)) > U^S(L_2(300))$  follows from the salience ranking,  $U^S(L_2(300)) = U^E(L_2(300))$  clearly holds and  $U^E(L_2(300)) > U^E(L_1(300))$  follows from diminishing sensitivity and risk aversion. Therefore, Bordalo *et al.* and our model predict that  $L_1(300)$  is chosen. For negative domains Bordalo *et al.* hypothesize that  $-L_2(300)$  will be chosen, whereas our model predicts  $-L_1(300)$ . The decision maker is optimistic that the relative downside of  $-L_1(300)$  will not materialize, so that she will go for the riskier option  $-L_1(300)$ . Thereby, she follows her utility function as she is risk-seeking in the negative domain.

### Prediction 3



- Under the assumption that  $s_1 = (1, 20)$  is more salient than  $s_2 = (381, 20)$ , our model and salience theory of choice under risk predict that subjects will choose  $L_2(0)$  and  $L_1(300)$ . For losses, salience theory predicts  $-L_1(0)$  and  $-L_2(300)$ , while our model predicts  $-L_1(0)$  and  $-L_1(300)$ .
- If  $s_1 = (1, 20)$  is less salient than  $s_2 = (381, 20)$ , our model and salience theory of choice under risk predict that subjects will choose  $L_1(0)$  and  $L_1(300)$ . For losses, salience theory predicts  $-L_2(0)$  and  $-L_2(300)$ , while our model predicts  $-L_2(0)$  and  $-L_1(300)$ .

As in our test for the common ratio Allais paradox, for gains our model’s predictions match those of salience theory. For losses, however, downsides are salient and overweighted according to (S), so that people choose the rational valuation (E), thereby making different choices than agents following Bordalo *et al.* (2012).

**Experimental findings.** Given  $s_1 = (1, 20)$  is more salient than  $s_2 = (381, 20)$ , then salience theory and our model both account for a share of 30% of the observations if we consider gains. This percentage is slightly lower than the 37% of subjects who decide in line with salience theory in the survey by Booth and Nolen (2013). Considering losses, salience theory accounts for a share of only 12.1%, while our model accounts for 31.3%.

Under the assumption that  $s_1 = (1, 20)$  is less salient than  $s_2 = (381, 20)$ , our model and salience theory account for 31.1% of the observations in the case of gains. Considering losses, our model performs significantly better than salience theory by correctly predicting a share of 36.1% compared to 20.5%.

	$L_1(0)$	$L_2(0)$			$-L_1(0)$	$-L_2(0)$	
$L_1(300)$	31.1	30	61.1	$-L_1(300)$	31.3	36.1	67.5
$L_2(300)$	8.9	30	38.9	$-L_2(300)$	12.1	20.5	32.5
	40.0	60	N = 90		43.4	56.6	N = 83

(a) Gains (choices in %)

(b) Losses (choices in %)

Table 3: Risk and Salience Attitudes

## 8 Conclusion

This paper proposes an attention-based theory of framing of intertemporal decision situations. The theory is based on two psychological phenomena: When making decisions, humans tend to overweight salient features. Further, humans tend to interpret information in a self-serving manner leading to our notion of an endogenously determined representation (frame) of a decision situation. By doing so we have also provided a unifying account of existing attention-based theories of decision making (Bordalo *et al.*, 2012; Köszegi and Szeidl, 2013).

As a main prediction we provide a rationale for overoptimistic choices that can often be observed in practice. The model gives an explanation as to how and why decision

makers underrate downside risks but overrate upside chances. In a single framework this can explain not only excessive gambling, but also excessive unhealthy diet. However, our model does not always predict overoptimistic choices. For instance, a risk-averse decision maker will not reveal overoptimistic choices when facing the choice between a safe option and a mean-preserving symmetric spread.

Our model may also prove useful for analyzing public policies. The setting where re-framing implies some psychological cost provides a framework of how informational policies may affect risk assessment and behavior. We hope that this framework of how to think about such informational policies and nudges might be helpful for researchers and policy makers in designing and evaluating such policy interventions.

Whereas we provide rather extreme examples in which external forces render certain aversive events salient, it remains for future work to investigate more precisely under which circumstances a decision maker can freely choose her mental choice representation and which external events pre-assign her to a certain frame. Therefore, future work could model or estimate switching costs between our proposed frames, possibly as a function of the strength of the external impact. For marketing and for policy issues, especially the further question of which interventions are more likely to pre-assign decision makers to certain frames or to enhance their mental switching costs between frames is interesting and remains for future research.

## References

- ANDREONI, J. and SPRENGER, C. (2010). Certain and uncertain utility: The Allais paradox and five decision theory phenomena. *Levine's Working Paper Archive*.
- BALCETIS, E. and DUNNING, D. (2006). See what you want to see: Motivational influence on visual perception. *Journal of Personality and Experimental Psychology*, **91** (4), 612–625.
- BANSAL-TRAVERS, M., HAMMOND, D., SMITH, P. and CUMMINGS, K. M. (2011). The impact of cigarette pack design, descriptors, and warning labels on risk perception in the US. *American Journal of Preventive Medicine*, **40** (6), 674–682.
- BATEMAN, I., MUNRO, A., RHODES, B., STARMER, C. and SUGDEN, R. (1997). A test of the theory of reference-dependent preferences. *Quarterly Journal of Economics*, **112** (2), 479–505.
- BAUER, T. K., BRAUN, S. and KVASNICKA, M. (2013). Distant event, local effects? Fukushima and the German housing market. *Ruhr Economic Papers* 433.
- BECKERT, J. and LUTTER, M. (2007). Wer spielt hat schon verloren. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, **59** (2), 240–270.
- BHATIA, S. and GOLMAN, R. (2013). Attention and reference dependence. *Working Paper*.
- BOOTH, A. and NOLEN, P. (2013). Salience, risky choices and gender. *Economics Letters*, **117**, 517–520.

- BORDALO, P., GENNAIOLI, N. and SHLEIFER, A. (2012). Salience theory of choice under risk. *Quarterly Journal of Economics*, **127** (3), 1243–1285.
- , — and — (2013). Salience and consumer choice. *Journal of Political Economy*, **121** (5), 803–843.
- BRUNNERMEIER, M. K. and PARKER, J. A. (2005). Optimal expectations. *American Economic Review*, **95** (4), 1092–1118.
- BUSHONG, B., RABIN, M. and SCHWARTZSTEIN, J. (2014). A model of relative thinking. *Working Paper*.
- CAMERER, C. and LOVALLO, D. (1999). Over-confidence and excess entry: An experimental approach. *American Economic Review*, **89**, 306–318.
- CUNNINGHAM, T. (2012). Comparisons and choice. *Working Paper*.
- DAWSON, E., GILOVICH, T. and REGAN, D. (2002). Motivated reasoning and performance on the Wason selection task. *Personality and Social Psychology Bulletin*, **28** (10), 1379–1387.
- DESSAINT, O. and MATRAY, A. (2014). Do managers overreact to salient risks? Evidence from hurricane strikes. *Working paper*.
- FALK, A. and ZIMMERMANN, F. (2014). Beliefs and utility: Experimental evidence on preferences for information. *Working Paper*.
- GABAIX, X. (2014). A sparsity-based model of bounded rationality. *Quarterly Journal of Economics*, **129** (4), 1661–1710.
- HADAR, J. and RUSSELL, W. R. (1969). Rules for ordering uncertain prospects. *American Economic Review*, pp. 25–34.
- HAMILTON, B. H. (2000). Does entrepreneurship pay? An empirical analysis of the returns of self-employment. *Journal of Political Economy*, **108** (3), 604–631.
- HASSAN, L. M., SHIU, E., THRASHER, J. F., FONG, G. T. and HASTINGS, G. (2008). Exploring the effectiveness of cigarette warning labels: Findings from the United States and United Kingdom arms of the international tobacco control (itc) four country survey. *International Journal of Nonprofit and Voluntary Sector Marketing*, **13** (3), 263–274.
- IPPOLITO, P. M. and IPPOLITO, R. A. (1984). Measuring the value of life saving from consumer reactions to new information. *Journal of Public Economics*, **25** (1-2), 53–81.
- KAHNEMAN, D. (2011). *Thinking, fast and slow*. Penguin.
- and TVERSKY, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, **47** (2), 263–291.

- KŐSZEGI, B. and RABIN, M. (2006). A model of reference-dependent preferences. *Quarterly Journal of Economics*, **121** (4), 1133–1165.
- KOELLINGER, P., MINNITI, M. and SCHADE, C. (2007). "I think I can, I think I can": Overconfidence and entrepreneurial behavior. *Journal of Economic Psychology*, **28**, 502–527.
- KŐSZEGI, B. and SZEIDL, A. (2013). A model of focusing in economic choice. *Quarterly Journal of Economics*, **128** (1), 53–104.
- LINDELL, M. K. and PERRY, R. W. (2000). Household adjustment to earthquake hazard: A review of research. *Environment and Behavior*, **32** (4), 461–501.
- MACHINA, M. J. (1989). Dynamic consistency and non-expected utility models of choice under uncertainty. *Journal of Economic Literature*, pp. 1622–1668.
- MALMENDIER, U. and TATE, G. (2005). Does overconfidence affect corporate investment? CEO overconfidence measures revisited. *European Financial Management*, **11** (5), 649–659.
- MAYRAZ, G. (2011). Priors and desires. *Working paper*.
- MESSICK, D. and SENTIS, K. (1979). Fairness and preference. *Journal of Experimental and Social Psychology*, **15** (4), 418–434.
- MOSKOVITZ, T. and VISSING-JORGENSEN, A. (2002). The returns to entrepreneurial investment: A private equity premium puzzle? *American Economic Review*, **92** (4), 745–778.
- MUNAFÒ, M. R., ROBERTS, N., BAULD, L. and LEONARDS, U. (2011). Plain packaging increases visual attention to health warnings on cigarette packs in non-smokers and weekly smokers but not daily smokers. *Addiction*, **106** (8), 1505–1510.
- OSTER, E., SHOULSON, I. and DORSEY, E. (2013). Limited life expectancy, human capital and health investments. *American Economic Review*, **103** (5), 1977–2002.
- PALM, R. (1995). Catastrophic earthquake insurance: Pattern of adoption. *Economic Geography*, **71** (2), 119–131.
- SAMUELSON, P. A. (1937). A note on measurement of utility. *Review of Economic Studies*, **4** (2), 155–161.
- SCHKADE, D. A. and KAHNEMANN, D. (1998). Does living in California make people happy? A focusing illusion in judgments of life satisfaction. *Psychological Science*, **9** (5), 340–346.
- SCHWARTZSTEIN, J. (2014). Selective attention and learning. *Journal of the European Economic Association*, **12** (6), 1423–1452.
- SHEPPERD, J. A., KLEIN, W. M. P., WATERS, E. A. and WEINSTEIN, N. D. (2013). Taking stock of unrealistic optimism. *Perspectives on Psychological Science*, **8** (4), 95–411.

- TAYLOR, S. and THOMPSON, S. (1982). Stalking the elusive vividness effect. *Psychological Review*, **89**, 155–181.
- TAYLOR, S. E. and BROWN, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, **103**, 193–210.
- THALER, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization*, **1**, 39–60.
- (1999). Mental accounting matters. *Journal of Behavioral Decision Making*, **12** (3), 183–206.
- WEINSTEIN, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, **39**, 806–820.
- (1989). Optimistic biases about personal risks. *Science*, **39**, 806–820.
- WOODFORD, M. (2012). Inattentive valuation and reference-dependent choice. *Working Paper*.

## Appendix

### Appendix A: Proof of Proposition 4

i) To show:  $U^S(A_1) > 0 \Rightarrow U^S(A_2) > 0$ . This holds true if  $U^S(A_2) - U^S(A_1) > 0$ . We show that this holds for the case that  $A_2$  is more positively skewed than  $A_1$  in a single period  $\tau$ , i.e., if  $|\mathcal{T}_1| = 1$ . A similar reasoning follows if  $|\mathcal{T}_1| > 1$ .

$U_r^S(A_2) - U_r^S(A_1) > 0 \Leftrightarrow \frac{p_\tau^2 g(G_\tau^2) G_\tau^2 - (1-p_\tau^2) g(L_\tau) L_\tau}{p_\tau^2 g(G_\tau^2) + (1-p_\tau^2) g(L_\tau)} - \frac{p_\tau^1 g(G_\tau^1) G_\tau^1 - (1-p_\tau^1) g(L_\tau) L_\tau}{p_\tau^1 g(G_\tau^1) + (1-p_\tau^1) g(L_\tau)} > 0 \Leftrightarrow [g(G_\tau^2) - g(G_\tau^1)] G_\tau^2 G_\tau^1 g(L_\tau) + [g(G_\tau^2) - g(L_\tau)] G_\tau^2 L_\tau g(G_\tau^1) + [g(L_\tau) - g(G_\tau^1)] G_\tau^1 L_\tau g(G_\tau^2) > 0$ , where the last step follows from using  $p_\tau^i = L_\tau / (L_\tau + G_\tau^i)$  and re-arranging. The first and second terms are positive, the third term negative. We consider two cases: a)  $G_\tau^2 g(G_\tau^1) > G_\tau^1 g(G_\tau^2)$  and b)  $G_\tau^2 g(G_\tau^1) < G_\tau^1 g(G_\tau^2)$

ad a) Re-arranging gives  $[g(G_\tau^2) - g(G_\tau^1)] G_\tau^2 G_\tau^1 g(L_\tau) + L_\tau \{ [g(G_\tau^2) - g(L_\tau)] G_\tau^2 g(G_\tau^1) + [g(L_\tau) - g(G_\tau^1)] G_\tau^1 g(G_\tau^2) \} > 0$ . The first term is positive. The second term is also positive as with  $G_\tau^2 g(G_\tau^1) > G_\tau^1 g(G_\tau^2)$  it holds that  $[g(G_\tau^2) - g(L_\tau)] G_\tau^2 g(G_\tau^1) + [g(L_\tau) - g(G_\tau^1)] G_\tau^1 g(G_\tau^2) > [g(G_\tau^2) - g(L_\tau)] G_\tau^2 g(G_\tau^1) + [g(L_\tau) - g(G_\tau^1)] G_\tau^2 g(G_\tau^1) = [g(G_\tau^2) - g(G_\tau^1)] G_\tau^2 g(G_\tau^1) > 0$ .

ad b) Re-arranging gives  $[g(G_\tau^2) - g(G_\tau^1)] G_\tau^2 G_\tau^1 g(L_\tau) + [G_\tau^2 - G_\tau^1] [g(G_\tau^1) g(G_\tau^2) L_\tau + [G_\tau^1 g(G_\tau^2) - G_\tau^2 g(G_\tau^1)] g(L_\tau) L_\tau] > 0$ , as all three terms are positive.

ii) To show:  $U^S(A_1) < 0 \Rightarrow U^S(A_2) < 0$  which holds if  $U^S(A_1) - U^S(A_2) > 0$ . As this can be shown via similar steps as under i) the details are omitted here.

### Appendix B: The Common ratio Allais paradox

We investigate under which conditions

$$U^S(L_1(0.001)) > U^S(L_2(0.002)) > U^E(L_2(0.002)) > U^E(L_1(0.001)).$$

and

$$U^S(L_2(0.9)) > U^S(L_1(0.45)).$$

hold. The Common Ratio Allais Paradox is consistent with salience theory for the wide range of salience parameters  $\delta \in (0.22, 1)$ , given Bordalo et al.'s salience function  $\sigma(x, y) = |x-y|/(|x|+|y|+0.1)$ . Both  $U^S(L_1(0.001)) > U^S(L_2(0.002))$  and  $U^S(L_2(0.9)) > U^S(L_1(0.45))$  hold for this parameter range. Therefore, in particular, these inequations can be assumed to hold also for our model.

The second inequality  $U^S(L_2(0.002)) > U^E(L_2(0.002))$  holds if focusing function  $g$ 's convexity is limited as follows. We impose the assumption

$$g(u(6000)) \leq 2g(u(3000)). \quad (\text{A1})$$

Formally,  $U^S(L_2(0.002)) > U^E(L_2(0.002))$  holds if the perceived winning probability exceeds the objective probability 0.002, i.e., if

$$\begin{aligned} 0.002 &< \frac{0.002 \cdot (0.998g(u(3000)) + 0.001g(u(6000) - u(3000)))}{0.001 \cdot (0.998g(u(6000)) + 0.002g(u(6000) - u(3000))) + 0.999 \cdot (0.002g(u(3000)) + 0.998g(0))} \\ &< \frac{0.002 \cdot 0.998g(u(3000))}{0.001 \cdot 0.998g(u(6000)) + 0.002 \cdot 0.998g(u(3000)) + 0.999 \cdot 0.998g(0)}. \end{aligned}$$

Given assumption (A1), this yields the very weak sufficient condition

$$g(u(3000)) > \frac{999}{997} g(0).$$

Therefore, given this lower boundary on the focusing function's steepness,  $L_2(0.002)$  yields a higher focus-weighted utility under a focus on states than under the rational valuation.

The third comparison  $U^E(L_2(0.002)) > U^E(L_1(0.001))$  always holds since we have incorporated a utility function satisfying diminishing sensitivity with respect to zero. Therefore, positive outcomes indicate that a decision maker is in her winning domain and consequently risk-averse. Then,  $2u(3000) > u(6000)$  so that she prefers the lower, more likely gain.