Second-best Probability Weighting*

Florian Herold
University of Bamberg

Nick Netzer
University of Zurich

November 2015

Abstract

Non-linear probability weighting is an integral part of descriptive theories of choice under risk such as prospect theory. But why do these objective errors in information processing exist? Should we try to help individuals overcome their mistake of overweighting small and underweighting large probabilities? In this paper, we argue that probability weighting can be seen as a compensation for preexisting biases in evaluating payoffs. In particular, inverse S-shaped probability weighting in prospect theory is a flipside of S-shaped payoff valuation. Probability distortions may thus have survived as a second-best solution to a fitness maximization problem, and it can be counterproductive to correct them while keeping the value function unchanged.

Keywords: probability weighting, prospect theory, evolution of preferences.

JEL Classification: D01, D03, D81.

---

*Email addresses: florian.herold@uni-bamberg.de and nick.netzer@econ.uzh.ch. We are grateful for very helpful comments by Orazio Attanasio, Adrian Bruhin, Robin Cubitt, Eddie Dekel, Charles Efferson, Jeff Ely, Thomas Epper, Daniel Friedman, Antonio Guarino, Aviad Heifetz, Steffen Huck, Philipp Kircher, Peter Klibanoff, Tatiana Kornienko, Christoph Kuzmics, Tim Netzer, Luis Rayo, Maria Saez-Marti, Joel Sobel, Philippe Tobler, seminar participants at the Universities of Aarhus, Bayreuth, Bielefeld, Cologne, Edinburgh, Heidelberg, Nottingham, Zurich, the 2011 meeting of the Verein für Socialpolitik in Frankfurt, the 2012 AEA meeting in Chicago, and the 2013 Workshop on Institutions, Individual Behavior and Economic Outcomes in Porto Conte. The paper also benefitted from conversations with Georg Nöldeke and Larry Samuelson about the general topic of evolutionary second-best. The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement PCIG11-GA-2012-322253. An earlier version of the paper was circulated under the title "Probability Weighting as Evolutionary Second-best".

1 Introduction

Prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992) is one of the most successful descriptive theories for choice under risk. It rests on two main building blocks for the evaluation of lotteries: an adaptive value function, which captures both loss aversion and different risk attitudes across the domains of gains and losses, and a probability weighting function, which captures systematic distortions in the way probabilities are perceived. Of these ingredients, the probability distortions are particularly remarkable, as they reflect objective errors in information processing: overweighting of small and underweighting of large probabilities, relative to their true magnitudes. The existence of such objective errors gives rise to puzzling questions. For instance, should we help individuals make better decisions by trying to correct their mistakes? Some scholars have argued that this is indeed the case.[1] Also, why do these errors exist in the first place? Evolutionary arguments typically predict that individuals who make systematic mistakes do not survive because they are replaced by more rational types.[2]

In this paper we propose a novel perspective on probability weighting, which sheds light on both its evolutionary origins and its consequences for paternalism. We argue that the particular form of probability misperception in prospect theory can be seen as an optimal compensation for similar biases introduced by the value function. Probability weighting may thus have survived as a second-best optimal solution to a fitness maximization problem, and it might be misleading to think of it as an error that needs correction.

Our basic model and results are presented in Section 2, where we consider the case of simple prospects with one possible payoff gain and one possible payoff loss. Nature offers such prospects randomly to an agent, who decides whether to accept or reject. The agent evaluates the payoffs by an S-shaped value function, as postulated in prospect theory. We investigate the agent's choices for different probability weighting schemes, and in particular we look for the shape of probability perception that yields choices with maximal expected payoffs. For the case without loss aversion, we first show that any solution to the problem indeed involves overweighting of small and underweighting of large probabilities. Intuitively, prospects with an expected payoff close to zero are especially prone to decision mistakes, and

[1] Camerer et al. (2003), for instance, argue in favor of a mild form of paternalism: "[s]ince low probabilities are so difficult to represent cognitively, it may help to use graphical devices, metaphors (imagine choosing a ping-poll ball out of a large swimming pool filled with balls), or relative odds comparisons (winning a lottery is about as likely as being struck by lightning in the next week)" (p. 1231).

[2] Robson and Samuelson (2010) survey a large literature on the evolution of preferences and behavior. More specifically, Robson (1996) has shown that evolution will select expected fitness maximizing agents whenever risk is idiosyncratic. Expected fitness maximization is not evolutionarily optimal for correlated risks (see also Cooper and Kaplan, 1982; Bergstrom, 1997; Curry, 2001; Robson and Samuelson, 2009), but correlation does not provide a rationale for the above described errors in probability perception.

small probabilities go along with large absolute payoffs in such prospects. Large absolute payoffs are, in turn, relatively undervalued by an S-shaped value function. To compensate, it becomes optimal to overweigh small and underweigh large probabilities. Non-linear probability weighting here emerges as a second-best distortion in response to the value function that constitutes a distortionary constraint (Lipsey and Lancaster, 1956). The resulting behavior is still different from expected payoff maximization in most cases, because the first-best cannot be achieved generically. We continue to show that the optimal perception can be implemented by reflective and symmetric weighting (Prelec, 1998) in some cases, and we discuss weighting functions that have been used in both empirical and theoretical work on prospect theory (Gonzalez and Wu, 1999). We also illustrate that loss aversion, the systematically different treatment of gains and losses, implies that gain and loss probabilities should be perceived differently, predicting violations of reflectivity in the direction of an optimism bias (Sharot et al., 2007). We finally discuss the inverse problem where the value function is chosen optimally given a non-linear probability weighting function. Section 3 of the paper contains affirming numerical results for more complex prospects.

Our results reveal an interesting internal structure of prospect theory by describing value and weighting function as two complementary elements that interact in an optimized way. This has implications for the interpretation of probability distortions, because now they serve a useful purpose rather than being a mistake. Related arguments about the expedience of biases have been made in the literature. Kahneman and Lovallo (1993) have first pointed out that exaggerated aversion to risk and exaggerated optimism might partially compensate for one another.[3] Besharov (2004) illustrates how attempts to correct interacting biases can backfire, focussing on overconfidence in conjunction with hyperbolic discounting and regret. The recent contribution by Steiner and Stewart (2014) derives distorted probability perception as an optimal correction for a winner's curse problem induced by noisy information processing. They also predict the prospect theory shape of probability weighting, and they discuss the opportunities and limits that this creates for "debiasing" agents.[4]

[3] Several other papers have provided purpose-oriented explanations for optimism or overconfidence. In Bernardo and Welch (2001), overconfidence helps to solve problems of herding. In Carrillo and Mariotti (2000), Bénabou and Tirole (2002) and Brocas and Carrillo (2004), individuals may choose to adhere to such delusions because they serve as commitment devices in the presence of time-inconsistency problems. Compte and Postlewaite (2004) model a situation where some overconfidence is optimal because it increases actual success probabilities. In Johnson and Fowler (2011), overconfidence arises in a hawk-dove type of interaction, based on the assumption that players behave aggressively whenever they believe to be more able than their opponent.

[4] In Friedman and Massaro (1998), agents cannot perceive or process a true objective probability $p \in [0, 1]$ but instead have to work with $\hat{p} \in [0, 1]$ which corresponds to the true probability p plus a noise term e. The signalling process is such that if $|p+e| > 1$ then a new signal is drawn, and thus the expected value of the true probability $E(p|\hat{p})$ is closer to 0.5 than the signal $\hat{p}$. Weighting probabilities is thus rational once the signal generating process is taken into account. Rieger (2014) also suggests a reason for probability weighting that

While evolutionary models quite often predict the emergence of fully rational agents, several contributions have modelled the effect of cognitive or perceptive constraints on the outcome of a biological selection process.[5] Related to prospect theory, some papers (e.g. Friedman, 1989; Robson, 2001; Rayo and Becker, 2007; Netzer, 2009) have concluded that adaptive and possibly S-shaped value functions can be superior in the presence of such constraints. In these models, steepness of the function that evaluates fitness payoffs can be helpful in preventing decision mistakes, because it enables to distinguish alternatives even if they are very similar to each other. A relatively large slope should thus be allocated to regions where correct decisions matter most, which can explain adaptation to a reference point and the S-shape.[6] In this paper, we treat the value function as a primitive and do not derive it from more basic principles. This can be thought of as a methodological shortcut that allows us to illustrate the interplay between payoff valuation and probability weighting without having to model a joint evolutionary process under constraints.

Some papers have explicitly investigated the interplay between different anomalies from an evolutionary perspective. With imperfect information constraints, Suzuki (2012) derives the joint evolutionary optimality of present bias, false beliefs and a concern to avoid cognitive dissonance, while Yao and Li (2013) show that optimism and loss aversion may coevolve because they jointly improve financial success, similar to one of our findings. Waldman (1994) was the first to apply the second-best concept in an evolutionary context. In his model a second-best population, consisting of agents who exhibit behavioral biases that are mutually (but not necessarily globally) optimal, cannot be invaded by a first-best mutant, because its optimal characteristics are diluted in the process of sexual recombination.[7] Ely (2011)

is different from ours. If observable by the opponent, the bias of underweighting large probabilities can have a strategic advantage in certain games such as war of attrition and a class of games he calls social control games. In such games it can favorably influence the mixed equilibrium and thus be evolutionary semi-stable. Heifetz et al. (2007) study the evolution of general perception biases that change the equilibrium structure in strategic settings. See Acemoglu and Yildiz (2001) for an evolutionary model of strategic interactions in which different behavioral anomalies emerge that perfectly compensate for one another.

[5] These include Samuelson (2004) and Noeldeke and Samuelson (2005), where agents cannot correctly process information about their environment, specifically about current earnings opportunities. Concern for relative consumption then becomes evolution's constrained optimal way of utilizing the information inherent in others' consumption levels. Samuelson and Swinkels (2006) argue that choice-set dependency might be an analogous way of correcting for an agent's lack of understanding of her choices' fitness implications under different conditions. Baliga and Ely (2011), although not explicitly within an evolutionary framework, start from the assumption of imperfect memory: after having made an investment, an agent forgets about the details of the project. The initial investment now still contains information, and relying on this information for future decisions on the project can be optimal rather than a "sunk cost fallacy".

[6] For related arguments in different modelling frameworks and for critical discussions see Robson and Samuelson (2011), Kornienko (2011), Wolpert and Leslie (2012) and Woodford (2012b,a). See Hamo and Heifetz (2002) for a model in which S-shaped utility evolves in a framework with aggregate risk, and McDermott et al. (2008) for a model where S-shaped utility is optimal if reproductive success is determined by a single fitness threshold.

[7] As an application, Waldman (1994) discusses the co-evolution of disutility of effort and overconfidence.

demonstrates that an evolutionary process can accumulate second-best "kludges" instead of ever converging to a first-best solution. He considers a complex species that remains stuck in a second-best forever because small local improvements, which take all other behavioral parameters as given, are substantially more likely to occur than a complete redesign. These general insights on the evolution of boundedly optimal behavior also justify our idea of optimizing the probability perception with a value function already in place (or the other way round), rather than solving a joint optimization problem. Even though it would be better to remove both distortions simultaneously, path-dependence of the evolutionary process can make such a solution very unlikely or infeasible.[8]

2 Probability Weighting for Simple Prospects

2.1 Basic Model

We start by considering a model of simple prospects. Such prospects consist of one possible payoff gain of size $x > 0$ which occurs with probability $0 < p < 1$, and one possible payoff loss of size $y > 0$ which occurs with probability $1 - p$. It is convenient to define prospects in terms of relative rather than absolute probabilities. Hence a simple prospect is a tuple (q, x, y) from the set $\mathscr{P} = \mathbb{R}^3_+$, where $q = p/(1 - p)$ is the relative probability of a gain. The function $F : \mathscr{P} \to \mathbb{R}$, defined by

$$F(q, x, y) = \left(\frac{q}{1+q}\right) x - \left(\frac{1}{1+q}\right) y, \tag{1}$$

assigns to each prospect (q, x, y) its expectation, where we used the identity $p = q/(1 + q)$. We will refer to F as the fitness function and treat it as the criterion for optimal decisions. With our evolutionary interpretation in mind, we think of gains and losses as being measured in terms of biological fitness relative to a decision-maker's current fitness level $c \in \mathbb{R}$ (such as the current number of expected offspring). The results by e.g. Robson (1996) then justify the use of (1) as the criterion for evolutionary success. In the simplest possible choice situation where the decision-maker is faced with a prospect and has to decide whether to accept it or

See Dobbs and Molho (1999) for a generalized version of the model, with an application to effort disutility and risk aversion, among others. Similar ideas are studied in Zhang (2013), who focusses on risk-aversion and overconfidence, and Frenkel et al. (2014), who examine the endowment effect and the winner's curse.

[8]For instance, we can imagine a payoff valuation function being deeply ingrained in our brains, something that may have evolved in our distant evolutionary past, shaped by evolutionary forces to deal with mostly simple and possibly deterministic choices that our brain was capable of analyzing at that time. Dealing with uncertainty in the form of explicit probabilities requires substantially more sophisticated cognitive capabilities, and the development of the corresponding cognitive processes might have taken place when the value function was already given.

stay at the current level of fitness with certainty, prospect (q, x, y) should be accepted if and only if $F(q, x, y) \geq 0$, or equivalently $q \geq y/x$, so that it (weakly) increases expected fitness above the current level. Let $\mathscr{P}^+ = \{(q, x, y) \in \mathscr{P} | q \geq y/x\}$ be the optimal acceptance set.

We now assume that the decision-maker uses an S-shaped value function $V : \mathbb{R} \to \mathbb{R}$ to evaluate gains and losses relative to c. It satisfies $V(c) = 0$ and is depicted in Figure 1. We decompose V into two functions, one used to evaluate gains (v_G) and one to evaluate losses (v_L). Specifically, we define $v_G : \mathbb{R}_+ \to \mathbb{R}_+$ by $v_G(x) = V(c + x)$ for all gains $x > 0$, and $v_L : \mathbb{R}_+ \to \mathbb{R}_+$ by $v_L(y) = -V(c - y)$ for all losses $y > 0$. We assume that both v_G and v_L are continuously differentiable, strictly increasing, strictly concave, and unbounded.

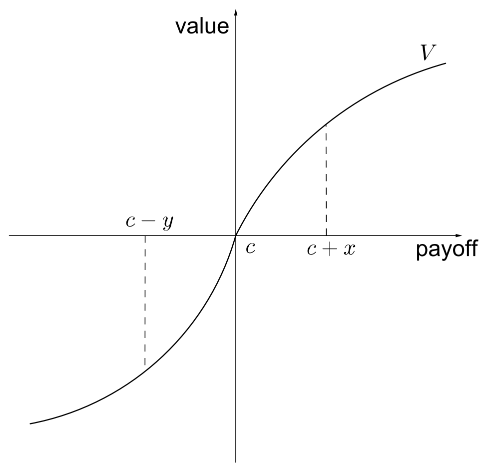

Figure 1: The value function

Concerning the perception of probabilities, let $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ be a measurable weighting function that yields a perceived relative probability $\eta(q)$ for any true relative probability q. Hence the actual gain probability p is perceived as $\eta(q)/(1 + \eta(q))$ and the loss probability $1 - p$ is perceived as $1/(1 + \eta(q))$ by the decision-maker. With these concepts at hand, the subjective utility score assigned to a prospect is

$$U_\eta(q, x, y) = \left(\frac{\eta(q)}{1 + \eta(q)}\right) v_G(x) - \left(\frac{1}{1 + \eta(q)}\right) v_L(y). \tag{2}$$

The decision-maker accepts prospect (q, x, y) if and only if $U_\eta(q, x, y) \geq 0$. This can be reformulated to $\eta(q) \geq v_L(y)/v_G(x)$, which is directly comparable to the criterion for optimal choice $q \geq y/x$. We will be interested in the weighting function η that maximizes expected fitness, given the fixed value function. Since the value function distorts one side of the optimality condition, it becomes apparent that the other side should be distorted as well. Let $\mathscr{P}^+_\eta = \{(q, x, y) \in \mathscr{P} | \eta(q) \geq v_L(y)/v_G(x)\}$ be the actual acceptance set.

We assume that nature randomly draws and offers to the decision-maker one prospect at a time, according to a probability distribution that can be described by a strictly positive density h on $\mathscr{P}$ (for which the maximal expected fitness $\bar{F}$, defined below, is finite). We are now interested in the solution to the following program, which we will also refer to as the fitness maximization problem:

$$\max_{\eta} \int_{\mathscr{P}_{\eta}^{+}} F(q,x,y)h(q,x,y)d(q,x,y). \tag{3}$$

Problem (3) can be thought of as the reduced form of a frequency-independent dynamic selection process (see e.g. Maynard Smith, 1978; Parker and Maynard Smith, 1990). Now suppose that (3) achieves the maximum of $\bar{F} = \int_{\mathscr{P}^{+}} F(q,x,y)h(q,x,y)d(q,x,y)$, i.e., there exists a weighting function η for which the actual and the optimal acceptance sets $\mathscr{P}^{+}$ and $\mathscr{P}_{\eta}^{+}$ coincide up to measure zero. Then behavior under η is almost surely identical to expected payoff maximization, so that η perfectly compensates for the value function. We will say that the first-best is achievable in this case. Otherwise, if the value of (3) is strictly smaller than $\bar{F}$, the solution is truly second-best, with observed behavior that deviates systematically from unconstrained expected payoff maximization.

The model of decision-making presented in this subsection corresponds to prospect theory as applied to simple prospects (see Section 2.5 for more details). It is, however, only a special case of a more general behavioral model that we present in the Appendix. This model does not rely on the utility specification (2), but only assumes that decision-making is driven by the comparison of two subjective values, the perceived relative probability $\eta(q)$ and a perceived relative payoff $W(x,y)$. Our following main results are corollaries of the more general results proven in the Appendix.

2.2 Optimal Weighting Without Loss Aversion

We first consider the case without loss aversion, which means that there is an S-shape but no systematic distortion in the perception of gains versus losses. Formally, the value function is symmetric and given by $v_G = v_L =: v$. The following proposition then states the important property that the relative probability of the gain is optimally overvalued if the gain is less likely than the loss, and undervalued otherwise.

Proposition 1. *Suppose there is no loss aversion. Then, any solution η^* to (3) satisfies*

$$\eta^*(q) \gtreqless q \quad \Leftrightarrow \quad q \lesseqgtr 1, \quad \text{for almost all } q \in \mathbb{R}_+. \tag{4}$$

To grasp an intuition for the result, assume that nature offers a prospect (q, x, y) where $q = y/x$, so that the prospect's expected payoff is exactly zero. If $q < 1$, then a decision-maker who perceives probabilities correctly would reject this prospect. The reason is that any such prospect must satisfy $x > y$, i.e., its gain must be larger than its loss, such that the S-shape of the value function (concavity of v) implies $v(y)/v(x) > y/x$, resulting in rejection due to an overvaluation of the loss relative to the gain. Monotonicity of v implies that all negative fitness prospects with $q < 1$ are then also rejected, but the same still holds for some positive fitness prospects, by continuity of v. Hence when there is no probability weighting, prospects with $q < 1$ are subject to only one of the two possible types of mistakes: rejection of prospects which should be accepted. The analogous argument applies to prospects with $q > 1$, where correct probability perception implies that only the mistake of accepting a prospect that should be rejected can occur. To counteract these mistakes, it becomes optimal to overvalue small and undervalue large relative probabilities. Note that this intuition does not depend on the specific density h. The exact shape of the solution η^* will generally depend on h, but the direction of probability weighting given in (4) does not.

2.3 Optimal Weighting With Loss Aversion

Next, we explore some implications of loss aversion for the optimal perception of probabilities. Loss aversion means that the decision-maker takes a loss of given size more seriously than a gain of the same size, so that the value function satisfies $v_G(z) < v_L(z)$ for all $z \in \mathbb{R}_+$. We work with a slightly stronger condition, which requires that there exists $\delta > 1$ such that $v_L(y)/v_G(x) > y/x$ whenever $y/x \leq \delta$. Hence the overvaluation of the loss relative to the gain must extend to some non-vanishing range where the loss is strictly larger than the gain. A simple and prominent special case is given by $v_G(x) = v(x)$ and $v_L(y) = \gamma v(y)$, for some common function v and a loss aversion parameter $\gamma > 1$, but there are also examples with non-multiplicative loss aversion (see the Appendix for details).

Proposition 2. *Suppose there is loss aversion. Then, there exists $\bar{q} > 1$ such that any solution η^* to (3) satisfies*

$$\eta^*(q) > q, \quad \text{for almost all } q \leq \bar{q}. \tag{5}$$

The result shows that loss aversion expands the range of overweighting, because correct probability perception would now lead to rejection of zero fitness prospects even when the gain is somewhat more likely (and thus smaller) than the loss. In particular, a gain should be perceived as more likely than the loss even when they are in fact equally probable ($\eta^*(1) > 1$).

Intuitively, if there is an asymmetry in the evaluation of gains and losses, the second-best principle calls for a systematically different treatment of gain and loss probabilities. We will discuss this in greater detail in Section 2.5.

2.4 First-best Solutions

How does the decision-maker's behavior look like if the probability perception is chosen to optimally compensate for the value function? We first ask the fundamental question whether optimized behavior is at all different from expected payoff maximization. Put differently, can the first-best be achieved despite the value function distortion, and if so, what is the first-best weighting function? We can give an answer to these questions as follows.

Proposition 3. *The first-best is achievable if and only if* $v_G(x) = \beta_G x^\alpha$ *and* $v_L(y) = \beta_L y^\alpha$ *for some* $\beta_G, \beta_L, \alpha > 0$. *In this case,* $\eta^{FB}(q) = (\beta_L/\beta_G)q^\alpha$ *is a first-best weighting function.*

Actual and optimal behavior must be aligned in a first-best solution. Specifically, all zero expected payoff prospects (q, x, qx) must be identified as such by the decision-maker. Formally, we need to ensure that $\eta(q) = v_L(qx)/v_G(x)$ always holds. This can be solved by some η if and only if the RHS $v_L(qx)/v_G(x)$ is independent of x, and the proposition provides the class of functions for which this is the case. Observe that this class is not generic: the first-best can only be achieved if the functions used to evaluate gains and losses are of the specific CRRA form, and loss aversion is compatible with the first-best only in multiplicative form with $\beta_G < \beta_L$. Otherwise, the solution η^* is truly second-best and the decision-maker's optimized behavior deviates from expected payoff maximization.

2.5 Relation to Prospect Theory

In prospect theory it is usually assumed that a function $\pi_G : [0,1] \to [0,1]$ transforms the gain probability p into a perceived gain decision weight $\pi_G(p)$ and a function $\pi_L : [0,1] \to [0,1]$ likewise transforms the loss probability $1-p$ into $\pi_L(1-p)$ (see e.g. Kahneman and Tversky, 1979; Prelec, 1998). The term decision weight is used because $\pi_G(p)$ and $\pi_L(1-p)$ do not necessarily sum up to one, and thus are not necessarily probabilities. The decision maker would then accept (q, x, y) if $\pi_G(q/(1+q))v_G(x) - \pi_L(1/(1+q))v_L(y) \geq 0$. Hence the absolute weighting functions π_G and π_L relate to our relative weighting function η as follows. On the one hand, any given pair (π_G, π_L) uniquely induces

$$\eta_{(\pi_G,\pi_L)}(q) = \frac{\pi_G(q/(1+q))}{\pi_L(1/(1+q))}.$$

The converse is not true, because there are different pairs of absolute weighting functions (π_G, π_L) that implement a given relative perception η. Hence there is not a unique way of representing the optimal weighting η^* in terms of prospect theory weighting functions, except if we impose additional requirements on (π_G, π_L).

One requirement that has received attention in both the theoretical and the empirical literature is reflectivity (Prelec, 1998). It postulates the identical treatment of gain and loss probabilities: $\pi_G(r) = \pi_L(r)$, $\forall r \in [0,1]$. We will skip the indices G and L when referring to reflective weighting functions. From an evolutionary perspective, reflective weighting is very appealing as it redundantizes the maintenance of two separate weighting mechanisms and thus saves on scarce cognitive resources. It is therefore plausible to assume that evolution would have settled on reflective weighting if the optimal probability perception can be implemented in this way. Even if no reflective weighting function implements our second-best solution, evolution may still favor reflective probability weighting if the savings in cognitive resources outweigh the losses from suboptimal decisions. Prelec (1998) summarizes evidence that indeed provides support for reflectivity. The following lemma identifies a condition that is necessary and sufficient for reflectivity to be no additional constraint in our framework.

Lemma 1. *A weighting function η can be implemented reflectively if and only if*

$$\eta(1/q) = 1/\eta(q), \quad \textit{for all } q \in \mathbb{R}_+. \tag{6}$$

Proof. Suppose η satisfies (6). Consider the candidate $\pi_\eta(p) = \eta(p/(1-p))/(1+\eta(p/(1-p)))$. It implements the relative perception

$$\frac{\pi_\eta\left(\frac{q}{1+q}\right)}{\pi_\eta\left(\frac{1}{1+q}\right)} = \frac{\eta\left(\frac{q}{1+q} \times \frac{1+q}{1}\right)}{1+\eta\left(\frac{q}{1+q} \times \frac{1+q}{1}\right)} \times \frac{1+\eta\left(\frac{1}{1+q} \times \frac{1+q}{q}\right)}{\eta\left(\frac{1}{1+q} \times \frac{1+q}{q}\right)} = \frac{\eta(q)}{1+\eta(q)} \frac{1+\eta(1/q)}{\eta(1/q)}.$$

Using (6) we obtain

$$\frac{\eta(q)}{1+\eta(q)} \frac{1+\eta(1/q)}{\eta(1/q)} = \frac{\eta(q)^2}{1+\eta(q)} \left[\frac{\eta(q)+1}{\eta(q)}\right] = \eta(q),$$

which shows that π_η indeed implements η reflectively. Conversely, suppose η is implemented reflectively by some π, i.e., $\eta(q) = \pi(q/(1+q)/\pi(1/(1+q))$ for all $q \in \mathbb{R}_+$. Then

$$\eta(1/q) = \frac{\pi\left(\frac{1/q}{1+1/q}\right)}{\pi\left(\frac{1}{1+1/q}\right)} = \frac{\pi\left(\frac{1}{1+q}\right)}{\pi\left(\frac{q}{1+q}\right)} = 1/\eta(q)$$

for all $q \in \mathbb{R}_+$, so η satisfies (6). □

Consider first the case without loss aversion and assume that the optimal η^* satisfies (6), in addition to the optimality condition (4). There is still not a unique reflective weighting function that implements η^*, but it does become unique when we impose the additional requirement of symmetry, which postulates $\pi(1-r) = 1-\pi(r)$, $\forall r \in [0,1]$ (see again Prelec, 1998). It is given by $\pi_{\eta^*}(p) = \eta^*(p/(1-p))/(1+\eta^*(p/(1-p)))$,[9] and straightforward calculations reveal that it satisfies

$$\pi_{\eta^*}(p) \gtreqless p \quad \Leftrightarrow \quad p \lesseqgtr \frac{1}{2}.$$

This corresponds to the key property of prospect theory that small probabilities are overweighted and large probabilities are underweighted. For instance, the specific function $\eta^{FB}(q) = (\beta_L/\beta_G)q^\alpha$ from Proposition 3 indeed satisfies (6) when $\beta_L = \beta_G$ and gives rise to

$$\pi_{\eta^{FB}}(p) = \frac{p^\alpha}{p^\alpha + (1-p)^\alpha},$$

which has been studied in the context of probability weighting by Karmarkar (1978, 1979). In Figure 2, it is depicted by a solid line for the case when $\alpha = 1/2$. However, there is also empirical evidence on the asymmetry of weighting functions. Individuals tend to overweigh absolute probabilities below approximately 1/3 and underweigh those above 1/3. Hence symmetry, while technically convenient, might actually not be a reasonable property to impose. Prelec (1998) instead suggests and axiomatizes the following reflective but non-symmetric function:

$$\pi(p) = e^{-[(-\ln p)^\alpha]} \quad \text{where} \quad 0 < \alpha < 1. \tag{7}$$

It is depicted by a dashed line in Figure 2, again for $\alpha = 1/2$. It can be shown that the relative perception η induced by this asymmetric weighting function also satisfies our necessary optimality condition (4).[10]

[9]Substituting $q = p/(1-p)$ in the condition that π_{η^*} implements η^* immediately implies the condition $\pi_{\eta^*}(p)/\pi_{\eta^*}(1-p) = \eta^*(p/(1-p))$. Using symmetry and rearranging then yields the given solution.

[10]Let η be implemented by a reflective π. Then $\eta(q) \gtreqless q$ can be rewritten as $\pi(p)/\pi(1-p) \gtreqless p/(1-p)$, which, for π given in (7), can be rearranged to

$$\left(\ln\left(\frac{1}{1-p}\right)\right)^\alpha - \left(\ln\left(\frac{1}{p}\right)\right)^\alpha + \ln\left(\frac{1}{p}\right) - \ln\left(\frac{1}{1-p}\right) \gtreqless 0.$$

Observe that the LHS of this inequality is indeed equal to 0 for $p = 1/2$ or $q = 1$, respectively. Our claim then follows from the fact that the LHS is strictly decreasing in p. The corresponding derivative is

$$\alpha\left(\frac{(-\ln(1-p))^{\alpha-1}}{1-p} + \frac{(-\ln(p))^{\alpha-1}}{p}\right) - \frac{1}{1-p} - \frac{1}{p}.$$

Its value is strictly increasing in α and for $\alpha = 1$ it is exactly 0. Hence it must be strictly negative for $\alpha < 1$.

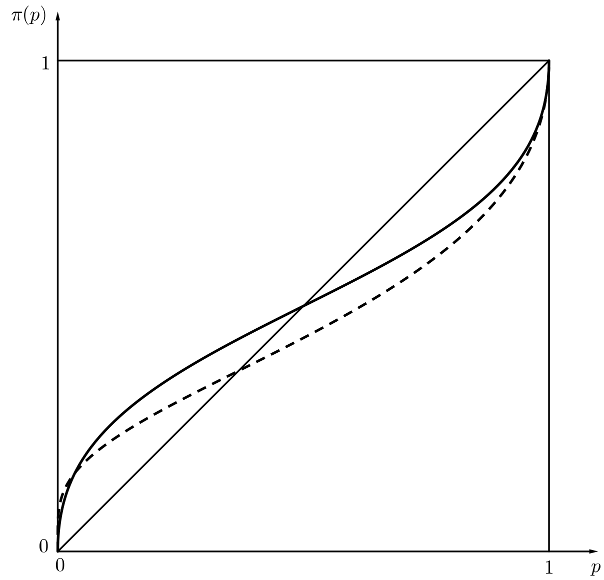

Figure 2: Common weighting functions

Consider now the case with loss-aversion, such that η^* satisfies condition (5). Any such η^* will violate (6), because we have $\eta^*(1/q) > 1/q > 1/\eta^*(q)$ for almost all $1 \leq q \leq \bar{q}$. The optimal absolute weighting functions must thus be non-reflective. As argued before, the wedge between the perception of gains and losses drives a second-best wedge between the perception of the associated probabilities. To understand the direction of this wedge, consider again $\eta^{FB}(q) = (\beta_L/\beta_G)q^\alpha$, now for $\beta_L > \beta_G$. The following "linear in log odds" functions (Gonzalez and Wu, 1999) implement η^{FB} and additionally satisfy $\pi_G^{FB}(p) + \pi_L^{FB}(1-p) = 1$:

$$\pi_G^{FB}(p) = \frac{(\beta_L/\beta_G)p^\alpha}{(\beta_L/\beta_G)p^\alpha + (1-p)^\alpha}, \qquad \pi_L^{FB}(p) = \frac{(\beta_G/\beta_L)p^\alpha}{(\beta_G/\beta_L)p^\alpha + (1-p)^\alpha}.$$

They are illustrated in Figure 3 for the case of $\alpha = 1/2$ and $\beta_L/\beta_G = 3/2$.[11] The function for gain probabilities lies above the one for loss probabilities. The model with loss aversion thus predicts a bias similar to overconfidence (Camerer and Lovallo, 1999) or, even more closely, optimism (Sharot et al., 2007). The fact that loss aversion and optimism might optimally compensate for one another has also been pointed out by Yao and Li (2013). They argue that investors who exhibit both anomalies might be most successful under certain circumstances.

[11] Several empirical studies have concluded that, to compensate for a loss of any given size, individuals require a gain of roughly twice that size (see e.g. Tom et al., 2007). With the functions $v_G(x) = \beta_G x^\alpha$ and $v_L(y) = \beta_L y^\alpha$ this is captured by $\beta_L/\beta_G = 2^\alpha$, which we approximate by 3/2 for the case when $\alpha = 1/2$.

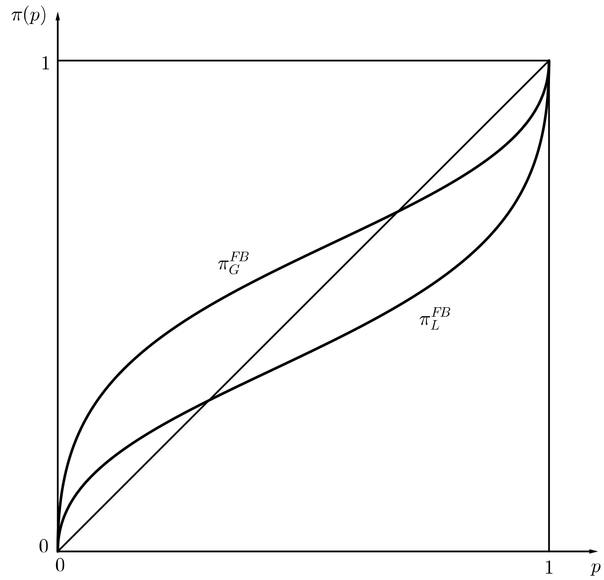

Figure 3: Non-reflective weighting for $\alpha = 1/2$ and $\beta_L/\beta_G = 3/2$.

2.6 The Inverse Problem

So far we have treated the value function as exogenous and the probability weighting function as endogenous. This has enabled us to shed some light on the question whether probability weighting is indeed a distortion in need of correction. It was also motivated in part by the earlier literature, which provided reasons for an S-shape of the value function without addressing the issue of probability weighting, and in part by the argument that the cognitive processes for dealing with explicit probabilities are presumably more advanced than simple valuation tasks and may thus have evolved in a second-best fashion later on. We can, however, also study the inverse problem of finding an optimal value function when a probability distortion is exogenously given. A formal treatment of this problem can be found in the Appendix. Here, we confine ourselves to a short discussion of the results.

If we capture inverse S-shaped probability weighting by property (4) for all $q \in \mathbb{R}_+$, then we can show in our generalized behavioral model that any optimally adapted payoff valuation function will exhibit an S-shape: the perception of the payoff ratio y/x will be enlarged (compressed) whenever the gain is larger (smaller) than the loss. This exactly mirrors our result for optimal probability weighting. Without further restrictions, however, we can always achieve the first-best in the inverse problem. The reason is that our generalized model allows for an independent perception $W(x, y)$ of any payoff pair (x, y), rather than working with the prospect theory inspired constraint that the ratio y/x has to be perceived as $v_L(y)/v_G(x)$ for two separate functions v_G and v_L. Once we impose this additional constraint, we retain a result analogous to Proposition 3: the first-best is achievable if and only if η is given by the specific functional form $\eta(q) = \gamma q^\alpha$.

3 Probability Weighting for More General Prospects

Working with simple prospects made the analysis tractable but poses the question to what extent our results are robust. Since models with more general prospects become hard to solve analytically, we present a numerical example in this section.

Assume that a prospect consists of a vector of payoffs $\mathbf{z} \in \mathbb{R}^n$, $n \geq 2$, where $\mathbf{z} = (z_1, ..., z_n)$ with $z_1 < ... < z_k < 0 < z_{k+1} < ... < z_n$ for $1 \leq k < n$, so that there are k possible losses and $n - k$ possible gains. The associated probabilities are given by $\mathbf{p} \in]0,1[^n$, where $\mathbf{p} = (p_1, ..., p_n)$ with $\sum_{i=1}^n p_i = 1$. The previous model is a special case for $n = 2$ and $k = 1$. The fitness mapping is given by $F(\mathbf{p}, \mathbf{z}) = \mathbf{p}\cdot\mathbf{z}$, where "$\cdot$" represents vector multiplication, and the optimal acceptance set $\mathscr{P}^+$ is defined as before. Let $\boldsymbol{\pi}$ be a weighting function that assigns decision weights $\boldsymbol{\pi}(\mathbf{p}) = (\pi_1(\mathbf{p}), ..., \pi_n(\mathbf{p}))$ to prospects. Assume that $U_{\boldsymbol{\pi}}(\mathbf{p}, \mathbf{z}) = \boldsymbol{\pi}(\mathbf{p}) \cdot V(\mathbf{z})$ is the decision-makers's utility from prospect $(\mathbf{p}, \mathbf{z})$, where $V(\mathbf{z}) \in \mathbb{R}^n$ denotes the vector obtained by mapping $\mathbf{z}$ pointwise through a value function $V : \mathbb{R} \to \mathbb{R}$. We then obtain the actual acceptance set $\mathscr{P}^+_{\boldsymbol{\pi}}$ of prospects with $U_{\boldsymbol{\pi}}(\mathbf{p}, \mathbf{z}) \geq 0$. The optimization problem we are interested in is given by

$$\max_{\boldsymbol{\pi}} \int_{\mathscr{P}^+_{\boldsymbol{\pi}}} F(\mathbf{p}, \mathbf{z}) h(\mathbf{p}, \mathbf{z}) d(\mathbf{p}, \mathbf{z}),$$

where $\boldsymbol{\pi}$ is chosen from some set of admissible functions.

Our following results are for two gains and two losses ($n = 4, k = 2$). We employ a value function given by $v_G(z) = z^\alpha$ and $v_L(z) = \gamma z^\alpha$, and we assume that nature offers prospects according to a uniform prior.[12] For each combination of $\alpha \in \{1, 3/4, 1/2, 1/4, 1/10\}$ and $\gamma \in \{1, 3/2\}$ we search for the optimal weighting function within a rank-dependent framework (Quiggin, 1982). There, the restriction is imposed that there exists a reflective

[12] Our numerical analysis is based on a finite model version. We discretize the interval $[0, 1]$ of probabilities into a grid of size n_p, i.e., we allow for probabilities $0, 1/n_p, 2/n_p, ..., 1$. We can then generate the set of all probability vectors $\mathbf{p} = (p_1, p_2, p_3, p_4)$ based on this grid. Analogously, we allow for payoffs between -2 and $+2$, discretized into a grid of size n_z. The set of all payoff vectors $\mathbf{z} = (z_1, z_2, z_3, z_4)$ can then be generated for that grid, and the set of all prospects is made up of all possible combinations of probabilities and payoffs. We use $n_p = n_z = 10$. The numerical calculations were performed in GNU Octave, and the script is available upon request.

function $w : [0,1] \to [0,1]$ which transforms cumulated probabilities.[13] Formally,

$$\pi_i(\mathbf{p}) = w\left(\sum_{j=1}^{i} p_i\right) - w\left(\sum_{j=1}^{i-1} p_i\right).$$

Table 1 contains the results when we use the function $w(p) = \delta p^\beta/(\delta p^\beta + (1-p)^\beta)$ and search for optimal values of β and δ in the range $[0,2]$, so that linear, S-shaped, and inverse S-shaped weighting is admitted.[14] The first main column refers to the case without loss aversion ($\gamma = 1$). If we additionally assume that payoffs are perceived linearly ($\alpha = 1$), the optimum does not involve any probability weighting ($\beta^* = \delta^* = 1$). The corresponding behavior is first-best, yielding the largest possible fitness level $\bar{F} = 0.2714$. We now introduce S-shaped payoff valuation by decreasing α towards zero. As the table shows, the optimal exponent of the weighting function then decreases, which means that probability weighting should become increasingly more inverse S-shaped. The maximum achievable fitness level also goes down as we increase the severity of the original distortion. Hence the first-best can no longer be achieved in the present setup with multiple gains and losses, despite the specific functional form of the value function that admitted a first-best solution for simple prospects. Without loss aversion the optimal weighting remains symmetric ($\delta^* = 1$). With loss aversion ($\gamma = 3/2$), the inverse S-shape of the weighting function is complemented by an asymmetry ($\delta^* < 1$). It is strikingly close to empirical findings, with the point of correct probability perception strictly below 1/2, as Figure 4 illustrates for the case where $\alpha = 1/2$ and $\gamma = 3/2$.

	$\gamma = 1$			$\gamma = 3/2$		
α	β^*	δ^*	Fitness	β^*	δ^*	Fitness
1	1.0000	1.0000	0.2714	1.0069	0.6944	0.2713
3/4	0.7708	1.0000	0.2713	0.7708	0.6875	0.2712
1/2	0.5208	1.0000	0.2707	0.5208	0.6806	0.2705
1/4	0.2639	1.0000	0.2690	0.2569	0.6736	0.2688
1/10	0.1042	1.0000	0.2672	0.1042	0.6667	0.2669

Table 1: $w(p) = \delta p^\beta/(\delta p^\beta + (1-p)^\beta)$.

[13]Cumulative prospect theory (Tversky and Kahneman, 1992) uses a similar formulation, where an inverse of cumulated probabilities is transformed for gains. We can obtain similar numerical results in this framework with non-reflective weighting.

[14]The search is carried out in two stages. First, we decompose $[0,2]$ into a grid of size $2n_g$ for $n_g = 12$ and identify the optimum on that grid. As a second step, we search the area around this optimum more closely, by again decomposing it into a grid of the same size.

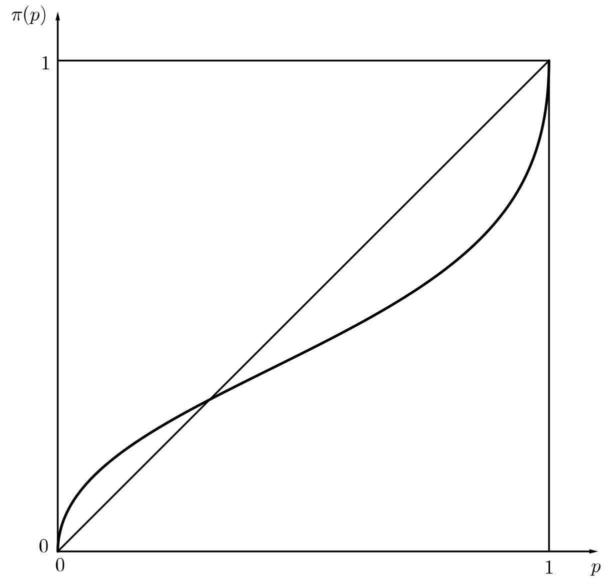

Figure 4: Numerical result for $\alpha = 1/2$, $\gamma = 3/2$.

4 Discussion

We conclude by discussing two avenues for future research. First, a caveat of our analysis is that either the value or the probability weighting function is treated as given, and the two are not derived jointly from a more basic constraint. A modelling framework such as the one recently proposed by Steiner and Stewart (2014) might be a promising candidate for a unified model. Based on an assumption about noisy information transmission, they can derive either an S-shaped value or an inverse S-shaped probability weighting function, but they do not have results about their interaction or joint optimality. Our own framework could also provide a starting point. For instance, consider the class of value and weighting functions that achieve the first-best (Proposition 3). In the spirit of the literature discussed in the Introduction, we could assume that steepness of the value and/or the probability weighting function is cognitively "expensive", which would provide an immediate argument for picking functions with $\alpha < 1$ out of the many behaviorally equivalent solutions.[15] A similar argument could still apply when the first-best is out of reach, e.g. for more complex prospects or when additional constraints have to be respected.

Second, since our analysis postulates that probability weighting is a complement to payoff valuation, it suggests that probability distortions are more pronounced for agents whose value function deviates more strongly from linearity. There seem to be few, if any, empirical

[15] We are grateful to Philipp Kircher for this suggestion.

studies that speak to this prediction about the correlation between the parameters of the two functions. Some studies have sample sizes that are too small to make meaningful correlation statements,[16] others do not address the question for different reasons.[17] Rieger et al. (2011) present estimated parameters for 45 countries, using the CRRA value function with different exponents for gains and losses, and a weighting function of the form

$$\pi(p) = p^\alpha / [p^\alpha + (1-p)^\alpha]^{1/\alpha}.$$

Based on their estimates (Table 2, p. 7) we obtain a correlation of about +0.23 between the weighting parameter and the gain exponent, and of about +0.02 for the loss exponent. These findings look promising but also suggest the need for additional empirical research.

References

Acemoglu, D. and Yildiz, M. (2001). Evolution of perception and play. Mimeo.

Baliga, S. and Ely, J. (2011). Mnemonomics: The sunk cost fallacy as a memory kludge. *American Economic Journal: Microeconomics*, forthcoming.

Bergstrom, T. (1997). Storage for good times and bad: Of rats and men. Mimeo.

Bernardo, A. and Welch, I. (2001). On the evolution of overconfidence and entrepreneurs. *Journal of Economics & Management Strategy*, 10:301–330.

Berns, G., Capra, C., Chappelow, J., Moore, S., and Noussair, C. (2008). Nonlinear neurobiological probability weighting functions for aversive outcomes. *NeuroImage*, 39:2047–2057.

Besharov, G. (2004). Second-best considerations in correcting cognitive biases. *Southern Economic Review*, 71:12–20.

[16] Gonzalez and Wu (1999) estimate parameters of the CRRA value function and the linear in log odds weighting function for 10 subjects. Using their results (Table 3, p. 157), we obtain a value of about −0.05 for the coefficient of correlation between the exponents of the value and the weighting function. Hsu et al. (2009) fit the CRRA value function together with the one parameter Prelec function (7) for 16 subjects. Based on their estimates (Table S4, p. 13 online supplementary material), we obtain a coefficient of correlation of about +0.04.

[17] Bruhin et al. (2010) estimate a finite mixture model in which 448 individuals are endogenously grouped into different behavioral categories. They find that two such categories emerge: 20% of all individuals maximize expected payoffs, using linear value and probability weighting functions, while 80% exhibit inverse S-shape probability weighting. However, Bruhin et al. (2010) find that the value function of the group that weighs probabilities non-linearly is not convex for losses. Qiu and Steiger (2011) estimate value and weighting functions for 124 subjects and conclude that there is no positive correlation. Their measure of the probability weighting function is, however, the relative area below the function, which captures elevation rather than the curvature property that we are interested in.

Bénabou, R. and Tirole, J. (2002). Self-confidence and personal motivation. *Quarterly Journal of Economics*, 117:871–915.

Brocas, I. and Carrillo, J. (2004). Entrepreneurial boldness and excessive investment. *Journal of Economics and Management Strategy*, 13:321–350.

Bruhin, A., Fehr-Duda, H., and Epper, T. (2010). Risk and rationality: Uncovering heterogeneity in probability distortion. *Econometrica*, 78:1375–1412.

Camerer, C., Issacharoff, S., Loewenstein, G., and O'Donoghue, T. (2003). Regulation for conservatives: Behavioral economics and the case for "asymmetric paternalism". *University of Pennsylvania Law Review*, 151:1211–1254.

Camerer, C. and Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review*, 89:306–318.

Carrillo, J. and Mariotti, T. (2000). Strategic ignorance as a self-disciplining device. *Review of Economic Studies*, 67:529–544.

Compte, O. and Postlewaite, A. (2004). Confidence-enhanced performance. *American Economic Review*, 94:1536–1557.

Cooper, S. and Kaplan, R. (1982). Adaptive "coin-flipping": A decision-theoretic examination of natural selection for random individual variation. *Journal of Theoretical Biology*, 94:135–151.

Curry, P. (2001). Decision making under uncertainty and the evolution of interdependent preferences. *Journal of Economic Theory*, 98:357–369.

Dobbs, I. and Molho, I. (1999). Evolution and sub-optimal behavior. *Journal of Evolutionary Economics*, 9:187–209.

Efthimiou, C. (2010). *Introduction to Functional Equations*. MSRI mathematical circles library.

Ely, J. (2011). Kludged. *American Economic Journal: Microeconomics*, 3:210–231.

Fehr-Duda, H., Bruhin, A., Epper, T., and Schubert, R. (2010). Rationality on the rise: Why relative risk aversion increases with stake size. *Journal of Risk and Uncertainty*, forthcoming.

Frenkel, S., Heller, Y., and Teper, R. (2014). The endowment effect as a blessing. Mimeo.

Friedman, D. (1989). The s-shaped value function as a constrained optimum. *American Economic Review*, 79:1243–1248.

Friedman, D. and Massaro, D. W. (1998). Understanding variability in binary and continuous choice. *Psychonomic Bulletin & Review*, 5 (3):370–389.

Gonzalez, R. and Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology*, 38:129–166.

Hamo, Y. and Heifetz, A. (2002). An evolutionary perspective on goal seeking and s-shaped utility. Mimeo.

Heifetz, A., Shannon, C., and Spiegel, Y. (2007). The dynamic evolution of preferences. *Economic Theory*, 32(2):251–286.

Hsu, M., Krajbich, I., Zhao, C., and Camerer, C. (2009). Neural Response to Reward Anticipation under Risk Is Nonlinear in Probabilities. *Journal of Neuroscience*, 18:2231–2237.

Johnson, D. and Fowler, J. (2011). The evolution of overconfidence. *Nature*, 477:317–320.

Kahneman, D. and Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science*, 39:17–31.

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

Karmarkar, U. (1978). Subjectively weighted utility: A descriptive extension of the expected utility model. *Organizational Behavior and Human Performance*, 21:61–72.

Karmarkar, U. (1979). Subjectively weighted utility and the allais paradox. *Organizational Behavior and Human Performance*, 24:67–72.

Kornienko, T. (2011). A cognitive basis for context-dependent utility. Mimeo.

Lipsey, R. and Lancaster, K. (1956). The general theory of second best. *Review of Economic Studies*, 24:11–32.

Maynard Smith, J. (1978). Optimization theory in evolution. *Annual Review of Ecology and Systematics*, 9:31–56.

McDermott, R., Fowler, J., and Smirnov, O. (2008). On the evolutionary origin of prospect theory preferences. *Journal of Politics*, 70:335–350.

Netzer, N. (2009). Evolution of time preferences and attitudes toward risk. *American Economic Review*, 99:937–955.

Noeldeke, G. and Samuelson, L. (2005). Information-based relative consumption effects: Correction. *Econometrica*, 73:1383–1387.

Parker, G. A. and Maynard Smith, J. (1990). Optimality theory in evolutionary biology. *Nature*, 348:27–33.

Platt, M. and Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature*, 400:233–238.

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

Qiu, J. and Steiger, E.-M. (2011). Understanding the two components of risk attitudes: An experimental analysis. *Management Science*, 57:193–199.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization*, 3:323–343.

Rayo, L. and Becker, G. (2007). Evolutionary efficiency and happiness. *Journal of Political Economy*, 115:302–337.

Rieger, M. (2014). Evolutionary stability of prospect theory preferences. *Journal of Mathematical Economics*, 50:1–11.

Rieger, M., Wang, M., and Hens, T. (2011). Prospect theory around the world. *SSRN Working Paper No. 1957606*.

Robson, A. (1996). A biological basis for expected and non-excpected utility. *Journal of Economic Theory*, 68:397–424.

Robson, A. (2001). The biological basis of economic behavior. *Journal of Economic Literature*, 39:11–33.

Robson, A. and Samuelson, L. (2009). The evolution of time preference with aggregate uncertainty. *American Economic Review*, 99:1925–1953.

Robson, A. and Samuelson, L. (2010). The evolutionary foundations of preferences. In Bisin, A. and Jackson, M., editors, *Handbook of Social Economics*, pages 221–310. North-Holland.

Robson, A. and Samuelson, L. (2011). The evolution of decision and experienced utilities. *Theoretical Economics*, 6:311–339.

Samuelson, L. (2004). Information-based relative consumption effects. *Econometrica*, 72:93–118.

Samuelson, L. and Swinkels, J. (2006). Information, evolution and utility. *Theoretical Economics*, 1:119–142.

Shafer, W. (1974). The nontransitive consumer. *Econometrica*, 42:913–919.

Sharot, T., Riccardi, A., Raio, C., and Phelps, E. (2007). Neural mechanisms mediating optimism bias. *Nature*, 450:102–196.

Steiner, J. and Stewart, C. (2014). Perceiving prospects properly. Mimeo.

Suzuki, T. (2012). Complementarity of behavioral biases. *Theory and Decision*, 72:413–430.

Tobler, P., Christopoulos, G., O'Doherty, J., Dolan, R., and Schultz, W. (2008). Neuronal distortions of reward probability without choice. *Journal of Neuroscience*, 28:11703–11711.

Tom, S., Fox, C., Trepel, C., and Poldrack, R. (2007). The neural basis of loss aversion in decision-making under risk. *Science*, 315:515–518.

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

Waldman, M. (1994). Systematic errors and the theory of natural selection. *American Economic Review*, 84:482–497.

Wolpert, D. and Leslie, D. (2012). Information theory and observational limitations in decision making. *The B.E. Journal of Theoretical Economics*, 12.

Woodford, M. (2012a). Inattentive valuation and reference-dependent choice. Mimeo.

Woodford, M. (2012b). Prospect theory as efficient perceptual distortion. *American Economic Review, Papers and Proceedings*, 102,3:41–46.

Yao, J. and Li, D. (2013). Bounded rationality as a source of loss aversion and optimism: A study of psychological adaptation under incomplete information. *Journal of Economic Dynamics & Control*, 37:18–31.

Zhang, H. (2013). Evolutionary justifications for non-bayesian beliefs. *Economics Letters*, 121:198–201.

Zhong, S., Israel, S., Xue, H., Ebstein, R., and Chew, S. (2009a). Monoamine oxidase a gene (maoa) associated with attitude towards longshot risks. *PLoS ONE*, 4(12):e8516.

Zhong, S., Israel, S., Xue, H., Sham, P., Ebstein, R., and Chew, S. (2009b). A neurochemical approach to valuation sensitivity over gains and losses. *Proceedings of the Royal Society B: Biological Sciences*, 276:4181–4188.

A General Approach to Optimal Probability Weighting

A.1 Generalized Model

As in the body of the paper, a prospect $(q, x, y) \in \mathscr{P} = \mathbb{R}^3_+$ consists of a gain $x > 0$, a loss $y > 0$, and a relative gain probability $q > 0$. The following assumption specifies the substance of our generalized decision-making model.

Assumption 1. (i) *The decision-maker accepts prospect* (q, x, y) *if and only if*

$$\eta(q) \geq W(x, y),$$

where $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ *is a measurable probability weighting function and* $W : \mathbb{R}^2_+ \to \mathbb{R}_+$ *is a payoff valuation function.*
(ii) $W(x, y)$ *is continuously differentiable with* $\partial W/\partial x < 0$ *and* $\partial W/\partial y > 0$*, and it satisfies* $\lim_{y\to\infty} W(x, y) = \infty$*,* $\lim_{y\to 0} W(x, y) = 0$*,* $\lim_{x\to\infty} W(x, y) = 0$*, and* $\lim_{x\to 0} W(x, y) = \infty$*.*

Remark 1. *In the body of the paper, a decision rule based on the utility function*

$$U_\eta(q, x, y) = \left(\frac{\eta(q)}{1 + \eta(q)}\right) v_G(x) - \left(\frac{1}{1 + \eta(q)}\right) v_L(y)$$

is considered, where v_i *(for* $i = G, L$*) is continuously differentiable, strictly increasing, strictly concave, and satisfies* $\lim_{z\to 0} v_i(z) = 0$*. The prospect* (q, x, y) *is accepted if* $U_\eta(q, x, y) \geq 0$*. This amounts to the case* $W(x, y) = v_L(y)/v_G(x)$*. The assumption that* $\lim_{y\to\infty} W(x, y) = \infty$ *and* $\lim_{x\to\infty} W(x, y) = 0$ *implies that* v_G *and* v_L *cannot be bounded. This assumption is made for simplicity only, and analogous results can be obtained in an extended model that allows for bounded value functions.*

Part (i) of Assumption 1 captures separability of perceived probabilities from perceived payoffs, i.e., the perception of q is independent from the perception of x and y.[18] Part (ii)

[18]It is an interesting question in how far the human brain processes probabilities separately from the evaluation of gains and losses. In an fMRI study, Berns et al. (2008) find behaviorally meaningful neural

requires that the acceptance threshold for the perceived relative gain probability should be continuously increasing in the amount of the potential loss and decreasing in the amount of the potential gain, hence capturing continuity and monotonicity.[19]

The fitness function $F : \mathscr{P} \to \mathbb{R}$ is given by

$$F(q,x,y) = \left(\frac{q}{1+q}\right)x - \left(\frac{1}{1+q}\right)y.$$

We denote by $\mathscr{P}^+ = \{(q,x,y) \in \mathscr{P} | q \geq y/x\}$ the subset of prospects that should be accepted in order to maximize expected fitness. Let $\mathscr{P}^+_\eta = \{(q,x,y) \in \mathscr{P} | \eta(q) \geq W(x,y)\}$ be the set of prospects that the decision-maker actually accepts. Prospects are drawn according to a strictly positive density h on $\mathscr{P}$ (for which $\bar{F}$ is finite), and thus the expected fitness maximization problem is given by

$$\max_\eta \int_{\mathscr{P}^+_\eta} F(q,x,y)h(q,x,y)d(q,x,y). \tag{8}$$

We investigate this problem under different assumptions in the following subsections.

A.2 Optimal Weighting Without Loss Aversion

We first add an assumption that captures the S-shape of the valuation function.

Assumption 2. *$W(x,y) \gtreqqless y/x$ if and only if $y/x \lesseqqgtr 1$.*

Assumption 2 requires a strictly enlarged (compressed) perception of the loss to gain ratio whenever the gain is strictly larger (smaller) than the loss. Its symmetry implies $W(z,z) = 1$ for all $z \in \mathbb{R}_+$ and therefore precludes loss aversion.

correlates of non-linear (inverse S-shape) probability weighting. They show that "the pattern of activation could be largely dissociated into magnitude-sensitive and probability-sensitive regions" (p. 2052), which they explicitly interpret as evidence for the hypothesis that "people process these two dimensions separately" (p. 2055). The only region that they found activated by both payoffs and probabilities is an area close to the anterior cingulate cortex, which they see as a "prime candidate for the integration of magnitude and probability information" (p. 2055). Note however that there also exists evidence of neural activity responding to both probabilities and payoffs, e.g. by Platt and Glimcher (1999) who study monkey neurons in the lateral intraparietal area. There are several other neuroscience studies that have also examined the neural basis of probability weighting. Tobler et al. (2008), for instance, investigate the coding of probability perception for non-choice situations. They also provide a literature review. See also Zhong et al. (2009a,b). Fehr-Duda et al. (2010) report an experimental instance of non-separability of payoffs and probabilities.

[19]After rewriting the decision rule in Assumption 1(i) as $\eta(q) - W(x,y) \geq 0$, it becomes reminiscent of the functional representation of a possibly non-transitive consumer in Shafer (1974). We are grateful to a referee for pointing this out.

Remark 2. *In the multiplicative utility model, Assumption 2 requires that* $v_G = v_L = v$ *and therefore* $W(x, y) = v(y)/v(x)$*. It is then satisfied by concavity of* v*.*

Proposition 4. *Under Assumptions 1 and 2, any solution* η^* *to (8) satisfies*

$$\eta^*(q) \gtreqless q \quad \Leftrightarrow \quad q \lesseqgtr 1, \quad \textit{for almost all } q \in \mathbb{R}_+.$$

Proof. Let Assumptions 1 and 2 be satisfied, and consider a prospect with relative probability q and loss y. How will a decision-maker behave who employs the functions η and W, depending on x? We can implicitly define a function $\tilde{x} : \mathbb{R}_+^2 \to \mathbb{R}_+$ by

$$W(\tilde{x}(y, r), y) = r$$

for all $y, r \in \mathbb{R}_+$. Assumption 1 implies that $\tilde{x}(y, r)$ is uniquely defined and continuously differentiable with $\partial\tilde{x}/\partial y > 0$ and $\partial\tilde{x}/\partial r < 0$ (by the implicit function theorem). The considered prospect will be accepted if and only if $x \geq \tilde{x}(y, \eta(q))$. Hence expected fitness can be written as

$$\int_{\mathbb{R}_+} \left[\int_{\mathbb{R}_+} \left[\int_{\tilde{x}(y,\eta(q))}^{\infty} F(q, x, y) h(q, x, y) dx \right] dy \right] dq.$$

For function η^* to maximize this expression, we need $\eta^*(q) \in \arg\max_{r \in \mathbb{R}_+} \Phi(r, q)$ for almost all $q \in \mathbb{R}_+$, where

$$\Phi(r, q) = \int_{\mathbb{R}_+} \left[\int_{\tilde{x}(y,r)}^{\infty} F(q, x, y) h(q, x, y) dx \right] dy.$$

We obtain the derivative

$$\frac{\partial\Phi(r, q)}{\partial r} = \int_{\mathbb{R}_+} \left[\left(-\frac{\partial\tilde{x}(y, r)}{\partial r} \right) F(q, \tilde{x}(y, r), y) h(q, \tilde{x}(y, r), y) \right] dy.$$

Both $-\partial\tilde{x}(y, r)/\partial r$ and $h(q, \tilde{x}(y, r), y)$ are strictly positive for all $q, y, r \in \mathbb{R}_+$. Now consider the remaining term $F(q, \tilde{x}(y, r), y)$, which has the same sign as $q\tilde{x}(y, r) - y$. We claim that it is strictly positive for all $y \in \mathbb{R}_+$ whenever $r \leq q < 1$. We have $W(\tilde{x}(y, r), y) = r$ by definition of $\tilde{x}$. When $r < 1$, Assumptions 1 and 2 then imply that $\tilde{x}(y, r) > y$, because $W(x, y)$ is strictly decreasing in x and $W(y, y) = 1$ must hold. Assumption 2 then further implies that $W(\tilde{x}(y, r), y) > y/\tilde{x}(y, r)$, which can be rearranged to $r\tilde{x}(y, r) - y > 0$. Since $\tilde{x}(y, r) > 0$ and $r \leq q$, this implies the claim $q\tilde{x}(y, r) - y > 0$. Hence we know that, whenever $q < 1$, we have $\partial\Phi(r, q)/\partial r > 0$ for any $r \leq q$. This implies that $\eta^*(q) > q$ must hold for almost all $q < 1$. Analogous arguments apply for the cases where $q = 1$ and $q > 1$. ☐

A.3 Optimal Weighting With Loss Aversion

To investigate the consequences of loss aversion, we replace Assumption 2 as follows.

Assumption 3. *There exists $\delta > 1$ such that $W(x, y) > y/x$ if $y/x \leq \delta$.*

Assumption 3 implies $W(z, z) > 1$ for all $z \in \mathbb{R}_+$, so that the decision-maker takes a loss of given size more seriously than a gain of the same size. More precisely, it requires a strictly enlarged perception of the loss to gain ratio even when the loss is larger than the gain, up to some bound $\delta > 1$ on the loss to gain ratio. The assumption is compatible with an S-shape, but the loss must be sufficiently larger than the gain ($y/x > \delta$) for a reversal of the relative distortion effect to occur.

Remark 3. *In the multiplicative utility model, Assumption 3 is satisfied for various different specifications. For instance, consider the case where $v_G(x) = v(x)$ and $v_L(y) = \gamma v(y)$ for a common function v and a parameter $\gamma > 1$ that measures the degree of loss aversion. For any $y/x \leq 1$ we have $W(x, y) = \gamma v(y)/v(x) > v(y)/v(x) \geq y/x$ by concavity of v. Now let the bound be $\delta = \gamma$. For any $1 < y/x \leq \delta$ we obtain $W(x, y) = \gamma v(y)/v(x) \geq (y/x)v(y)/v(x) > y/x$, so that Assumption 3 is satisfied. As another example, consider the case where $v_G(x) = (x + 1)^\beta - 1$ and $v_L(y) = (y + 1)^\alpha - 1$ for $0 < \beta < \alpha < 1$ (adding and subtracting 1 ensures $v_G(z) < v_L(z)$ for all $z \in \mathbb{R}_+$). For any $y/x \leq 1$ we have $W(x, y) = ((y + 1)^\alpha - 1)/((x + 1)^\beta - 1) > ((y + 1)^\beta - 1)/((x + 1)^\beta - 1) \geq y/x$ due to $0 < \beta < \alpha < 1$. Now set the bound $\delta = \alpha/\beta$. For any $1 < y/x \leq \delta$ we obtain $W(x, y) = ((y + 1)^\alpha - 1)/((x + 1)^\beta - 1) > ((x + 1)^\alpha - 1)/((x + 1)^\beta - 1)$. The last term can be shown to be strictly larger than α/β (it converges to α/β as $x \to 0$). Hence $W(x, y) > \alpha/\beta \geq y/x$, which again shows that Assumption 3 is satisfied.*

Proposition 5. *Under Assumptions 1 and 3, there exists $\bar{q} > 1$ such that any solution η^* to (8) satisfies*

$$\eta^*(q) > q, \quad \text{for almost all } q \leq \bar{q}.$$

Proof. The proof follows the one for Proposition 4, up to the point where we now need to show that there exists $\bar{q} > 1$ such that $q\tilde{x}(y, r) - y$ is strictly positive for all $y \in \mathbb{R}_+$ whenever $r \leq q \leq \bar{q}$. Set $\bar{q} = \delta$ for a bound $\delta > 1$ as described in Assumption 3. We have $W(\tilde{x}(y, r), y) = r$ by definition of $\tilde{x}$. When $r \leq \bar{q} = \delta$, Assumptions 1 and 3 then imply that $\tilde{x}(y, r) > y/\delta$, because $W(x, y)$ is strictly decreasing in x and $W(y/\delta, y) > \delta$ holds. Assumption 3 then further implies that $W(\tilde{x}(y, r), y) > y/\tilde{x}(y, r)$, which can be rearranged to $r\tilde{x}(y, r) - y > 0$. Since $\tilde{x}(y, r) > 0$ and $r \leq q$, this implies the claim $q\tilde{x}(y, r) - y > 0$. The rest of the argument is again as in the proof of Proposition 4. □

A.4 First-Best Solutions

We now investigate the assumptions under which the first-best is achievable, i.e., under which a solution η^* to (8) induces a set $\mathscr{P}^+_{\eta^*}$ that coincides with $\mathscr{P}^+$ up to measure zero.

Proposition 6. *Under Assumption 1, the first-best is achievable if and only if there exists a function $f : \mathbb{R}_+ \to \mathbb{R}_+$ such that $W(x,y) = f(y/x)$. Then, $\eta^{FB} = f$ is a first-best solution.*

Proof. Step 1. Suppose that the first-best is achieved by η^{FB}. Using function $\tilde{x}$ defined in the proof of Proposition 4, we then have that $\mathscr{P}^+_{\eta^{FB}} = \{(q,x,y) \in \mathscr{P} | x \geq \tilde{x}(y, \eta^{FB}(q))\}$ and $\mathscr{P}^+ = \{(q,x,y) \in \mathscr{P} | x \geq y/q\}$ coincide up to measure zero. This implies that, for almost all $q \in \mathbb{R}_+$, we have $\tilde{x}(y, \eta^{FB}(q)) = y/q$ for almost all $y \in \mathbb{R}_+$, and hence for all $y \in \mathbb{R}_+$ by continuity of $\tilde{x}$. Substituting this into the equation that defines $\tilde{x}$ we obtain that, for almost all $q \in \mathbb{R}_+$, $W(y/q, y) = \eta^{FB}(q)$ holds for all $y \in \mathbb{R}_+$. Continuity of W then implies that there exists a function $f : \mathbb{R}_+ \to \mathbb{R}_+$ (which coincides with η^{FB} almost everywhere) such that $W(y/q, y) = f(q)$ for all $q, y \in \mathbb{R}_+$. Now consider any pair $x, y \in \mathbb{R}_+$. We obtain $W(x,y) = W(y/(y/x), y) = f(y/x)$.

Step 2. Suppose that there exists a function $f : \mathbb{R}_+ \to \mathbb{R}_+$ such that $W(x,y) = f(y/x)$. Under Assumption 1, f must be strictly increasing. With weighting function $\eta^{FB}(q) = f(q)$ we then obtain $\eta^{FB}(q) \gtreqqless W(x,y) \Leftrightarrow f(q) \gtreqqless f(y/x) \Leftrightarrow q \gtreqqless y/x$, so that $\mathscr{P}^+_{\eta^{FB}}$ and $\mathscr{P}^+$ coincide and the first-best is achieved. □

This result does not rely on Assumption 2 (S-shape) or Assumption 3 (loss aversion). For instance, without any payoff distortion ($W(x,y) = y/x$) the first-best is obviously achievable by correct probability perception ($\eta^{FB}(q) = q$). However, there are also valuation functions with an S-shape that satisfy the condition stated in Proposition 6: $W(x,y) = f(y/x)$ satisfies Assumption 2 when $f(z) \gtreqqless z$ if and only if $z \lesseqqgtr 1$. Similarly, there are valuation functions with loss aversion that satisfy the condition stated in Proposition 6: Assumption 3 holds when $f(z) > z$ for all $z \leq \delta$. While neither S-shape nor loss aversion are an impediment to attainability of the first-best in principle, the following remark illustrates that first-best solutions can still not be considered a generic case.

Remark 4. *In the multiplicative model with $W(x,y) = v_L(y)/v_G(x)$, the first-best is achievable according to Proposition 6 if and only if $v_L(y)/v_G(x) = f(y/x)$ holds for all $x, y \in \mathbb{R}_+$. It follows from Lemma 2 in Appendix C that this is the case if and only if $v_G(x) = \beta_G x^\alpha$ and $v_L(y) = \beta_L y^\alpha$ for $\beta_G, \beta_L, \alpha > 0$ (the lemma is applicable after substituting $a = x$, $b = y/x$, and relabelling $f_1 = v_G$, $f_2 = f$, $f_3 = v_L$). In this case, $\eta^{FB}(q) = (\beta_L/\beta_G)q^\alpha$ is a first-best weighting function. Hence the first-best can only be achieved if the functions used to evaluate* *gains and losses are of the specific CRRA form. In addition, loss aversion is compatible with the first-best only in multiplicative form with* $\beta_G < \beta_L$.

B General Approach to Optimal Payoff Valuation

We now consider the inverse problem where a weighting function η is given with properties as described in Proposition 4, and we derive properties of an optimally adapted valuation function W. We proceed in parallel to the previous section but restate the necessary concepts.

Assumption 1'. (*i*) *The decision-maker accepts prospect* (q, x, y) *if and only if*

$$\eta(q) \geq W(x, y),$$

where $W : \mathbb{R}_+^2 \to \mathbb{R}_+$ *is a measurable payoff valuation function and* $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ *is a probability weighting function.*
(*ii*) $\eta(q)$ *is continuously differentiable with* $\partial\eta/\partial q > 0$*, and it satisfies* $\lim_{q\to\infty} \eta(q) = \infty$ *and* $\lim_{q\to 0} \eta(q) = 0$.

Given Assumption 1', let $\mathscr{P}_W^+ = \{(q, x, y) \in \mathscr{P} | \eta(q) \geq W(x, y)\}$ be the set of prospects that are accepted. The expected fitness maximization problem is then given by

$$\max_W \int_{\mathscr{P}_W^+} F(q, x, y)h(q, x, y)d(q, x, y). \tag{9}$$

We next add an assumption that captures the inverse S-shape of the weighting function.

Assumption 2'. $\eta(q) \gtreqless q$ *if and only if* $q \lesseqgtr 1$.

Proposition 7. *Under Assumptions 1' and 2', any solution* W^* *to (9) satisfies*

$$W^*(x, y) \gtreqless y/x \;\;\Leftrightarrow\;\; y/x \lesseqgtr 1, \;\; \text{for almost all } (x, y) \in \mathbb{R}_+^2.$$

Proof. Let Assumptions 1' and 2' be satisfied, and consider a prospect with gain x and loss y. A decision-maker who employs η and W will accept the prospect if and only if

$$q \geq \eta^{-1}\left(W(x, y)\right),$$

where Assumption 1' implies that the inverse function $\eta^{-1}(r)$ of $\eta(q)$ is well-defined and continuously differentiable with $\partial\eta^{-1}/\partial r > 0$. Hence expected fitness can be written as

$$\int_{\mathbb{R}_+^2}\left[\int_{\eta^{-1}(W(x,y))}^{\infty} F(q,x,y)h(q,x,y)dq\right] d(x,y).$$

For function W^* to maximize this expression, we need $W^*(x,y) \in \arg\max_{r\in\mathbb{R}_+} \Phi(r,x,y)$ for almost all $(x,y) \in \mathbb{R}_+^2$, where

$$\Phi(r,x,y) = \int_{\eta^{-1}(r)}^{\infty} F(q,x,y)h(q,x,y)dq.$$

We obtain the derivative

$$\frac{\partial\Phi(r,x,y)}{\partial r} = -\left(\frac{\partial\eta^{-1}(r)}{\partial r}\right) F(\eta^{-1}(r),x,y)h(\eta^{-1}(r),x,y).$$

Both $\partial\eta^{-1}(r)/\partial r$ and $h(\eta^{-1}(r),x,y)$ are strictly positive for all $x,y,r \in \mathbb{R}_+$. Now consider the remaining term $F(\eta^{-1}(r),x,y)$, which has the same sign as $\eta^{-1}(r)x - y$. We claim that it is strictly negative whenever $r \leq y/x < 1$. When $r < 1$, Assumption 2' implies that $\eta^{-1}(r) < r$. Hence $\eta^{-1}(r)x - y < rx - y \leq (y/x)x - y = 0$, which establishes the claim. Hence we know that, whenever $y/x < 1$, we have $\partial\Phi(r,x,y)/\partial r > 0$ for any $r \leq y/x$. This implies that $W^*(x,y) > y/x$ must hold for almost all (x,y) with $y/x < 1$. Analogous arguments apply for the cases where $y/x = 1$ and $y/x > 1$. □

For ease of comparison, we have stated and proven Proposition 7 in exactly the same way as Proposition 4, but here we can additionally solve the first-order condition derived in the proof to obtain the optimal payoff valuation function $W^*(x,y) = \eta(y/x)$. It follows immediately (e.g. from Proposition 6) that this solution is always first-best. The reason is of course that here we are free to independently design the perception of any pair (x,y), while only the perception of q could be designed in the previous section. Let us therefore consider the additional constraint that $W(x,y) = v_L(y)/v_G(x)$ must hold, and only the (continuously differentiable and strictly increasing) functions v_L and v_G can be chosen.

Proposition 8. *Under Assumption 1' and the constraint that* $W(x,y) = v_L(y)/v_G(x)$*, the first-best is achievable if and only if* $\eta(q) = \gamma q^\alpha$ *for* $\gamma,\alpha > 0$*. Then, any pair* $v_G^{FB}(x) = \beta_G x^\alpha$ *and* $v_L^{FB}(y) = \beta_L y^\alpha$ *with* $\beta_L/\beta_G = \gamma$ *is a first-best solution.*

Proof. Step 1. Suppose that the first-best is achieved by $W^{FB}(x,y) = v_L^{FB}(y)/v_G^{FB}(x)$. Using the function η^{-1} as defined in the proof of Proposition 7, we then have that $\mathscr{P}^+_{W^{FB}} = \{(q,x,y) \in \mathscr{P} | q \geq \eta^{-1}(v_L^{FB}(y)/v_G^{FB}(x))\}$ and $\mathscr{P}^+ = \{(q,x,y) \in \mathscr{P} | q \geq y/x\}$ coincide up to measure zero. This implies that, for almost all $(x,y) \in \mathbb{R}_+^2$, we have $v_L^{FB}(y)/v_G^{FB}(x) = \eta(y/x)$.

Continuity of v_G^{FB}, v_L^{FB} and η in fact implies that the equality holds for all $(x, y) \in \mathbb{R}_+^2$. Then it follows from Lemma 2 in Appendix C that η must be of the form $\eta(q) = \gamma q^\alpha$ for $\gamma, \alpha > 0$ (the lemma is applicable after substituting $a = x$, $b = y/x$, and relabelling $f_1 = v_G^{FB}, f_2 = \eta, f_3 = v_L^{FB}$).

Step 2. Suppose that $\eta(q) = \gamma q^\alpha$ for $\gamma, \alpha > 0$. For any pair $v_G^{FB}(x) = \beta_G x^\alpha$ and $v_L^{FB}(y) = \beta_L y^\alpha$ with $\beta_L/\beta_G = \gamma$ we then obtain $\eta(q) \gtreqless v_L^{FB}(y)/v_G^{FB}(x) \Leftrightarrow \gamma q^\alpha \gtreqless \gamma(y/x)^\alpha \Leftrightarrow q \gtreqless y/x$, so that $\mathscr{P}^+_{W^{FB}}$ and $\mathscr{P}^+$ coincide and the first-best is achieved. □

C Multiplicative Functions

The following lemma is useful for the characterization of first-best solutions in the multiplicative utility model. The result is known in the literature on functional equations as the solution of the fourth Pexider functional equation (or power Pexider functional equation).

Lemma 2. *Three continuous functions $f_i : \mathbb{R}_+ \to \mathbb{R}, i = 1, 2, 3$, satisfy the functional relation*

$$f_1(a)f_2(b) = f_3(ab) \quad \textit{for all} \quad a, b \in \mathbb{R}_+ \tag{10}$$

if and only if

$$f_1(x) = \beta_1 x^\alpha,\ f_2(x) = \beta_2 x^\alpha,\ f_3(x) = \beta_1\beta_2 x^\alpha \quad \textit{for some} \quad \alpha, \beta_1, \beta_2 \geq 0. \tag{11}$$

A proof (of a slightly more general version) can e.g. be found in Section 8.4 of Efthimiou (2010, p. 142). The proof uses the solution to the power Cauchy equation, which is derived in Section 5.4 of the same book (p. 97f).