

Searching for the Reference Point

AURÉLIEN BAILLON
Erasmus School of Economics

HAN BLEICHRODT
Erasmus School of Economics

VITALIE SPINU
Erasmus School of Economics

8 December 2015

Abstract

This paper presents evidence on the formation of reference points. In a high-stakes experiment with payoffs up to a weekly salary, we found that most subjects used the status quo or a security-level (the maximum of the minimal outcomes of the prospects under consideration) as their reference point. Between ten and twenty percent of the subjects used expectations-based reference point as in the model of Köszegi and Rabin (2006, 2007).

Key words: formation reference points, reference-dependence, Bayesian hierarchical modeling.

1. Introduction

A key insight of behavioral economics is that people evaluate outcomes as gains and losses from a reference point. Reference-dependence is central in prospect theory, currently the most influential theory of decision under risk, and it plays a crucial role in explaining people's attitudes towards risk (Rabin, 2000). A lot of evidence, from both the lab and the field, supports reference-dependent preferences.¹

Prospect theory and other reference-dependent theories are typically silent about how reference points are formed. Tversky and Kahneman (1991) argued that "although the reference point usually corresponds to the decision maker's current position it can also be influenced by aspirations, expectations, norms, and social comparisons." This silence is undesirable as it creates too much freedom in deriving predictions. If the reference point is a free variable, almost any observed behavior can be explained and it becomes hard, if not impossible, to test reference-dependent theories empirically. For example, as Pesendorfer (2006) points out, different assumptions about the reference point have to be made to explain two well-known anomalies from finance: the equity premium puzzle demands that the reference point adjusts over time, whereas the disposition effect demands that the reference point remains constant at the purchase price. Markowitz (1952) already pointed out the importance of finding the reference point.² In his recent review of the literature, more than 60 years after Markowitz,

¹ Examples of real-world evidence for reference-dependence are the equity premium puzzle, the finding that stock returns are too high relative to bond returns (Benartzi and Thaler, 1995), the disposition effect, the finding that investors hold losing stocks and property too long and sell winners too early (Odean, 1998, Genesove and Mayer, 2001), default bias in pension and insurance choice (Samuelson and Zeckhauser, 1988, Thaler and Benartzi, 2004) and organ donation (Johnson and Goldstein, 2003), the excessive buying of insurance (Sydnor, 2010), the annuitization puzzle, the fact that at retirement people allocate too little of their wealth to annuities (Benartzi et al., 2011), the behavior of professional golf players (Pope and Schweitzer, 2011) and poker players (Eil and Lien, 2014), and the bunching of marathon finishing times just ahead of round numbers (Allen et al., 2013).

² In Markowitz's (1952) approach, the "customary wealth", which may deviate from present wealth, plays the role of a reference point. Markowitz (1952) said: "It would be convenient if I had a formula from which

Barberis (2013) still concludes that addressing the formation of the reference point is a key challenge to apply prospect theory to economics (p.192).

Several theories of reference point formation have been put forward. Heath et al. (1999) suggested that people's goals serve as their reference points. Köszegi and Rabin (2006, 2007) proposed a model in which the reference point is based on people's (rational) expectations. Their model is close in spirit to the disappointment models of Bell (1985), Loomes and Sugden (1986), Gul (1991), and Delquié and Cillo (2006) in which decision makers also form expectations about uncertain prospects and experience elation or disappointment depending on whether the actual outcome is better or worse than those expectations. Diecidue and Van de Ven (2008) presented a model with an aspiration level, which is a form of reference dependence.

Empirical evidence on the formation of reference points is scarce. Some evidence is consistent with Köszegi and Rabin's model of expectations-based reference points (Abeler et al., 2011, Crawford and Meng, 2011, Card and Dahl, 2011, Gill and Prowse, 2012, Bartling et al., 2015). On the other hand, the data in Baucells et al. (2011), Allen et al. (2013), and Lien and Zheng (2015) are inconsistent with Köszegi and Rabin's model and Barberis (2013) concludes that in finance there are "natural reference points other than expectations." Evidence from medical decision making suggests that, instead of an expectations-based reference point, people adopt a security-based reference point: the minimum health state they can reach for sure (van Osch et al., 2004, van Osch et al., 2006, Bleichrodt et al., 2001).

customary wealth could be calculated when this was not equal to present wealth. But I do not have such a rule and formula. For some clear-cut cases I am willing to assert that there are or are not recent windfall gains or losses: the man who just won or lost at cards; the man who has experienced no change in income for years. I leave it to the reader's intuition to recognize other clear-cut cases. I leave it to future research and reflection to classify the ambiguous, border-line cases."

The purpose of this paper is to shed new light on the formation of reference points in decisions under uncertainty. The reference points we consider can all be identified through choices and, hence, we work within the revealed preference paradigm. These reference points can also be easily applied in practical research and require no prior knowledge. Moreover, we do not use any other input than what is traditionally used in decision theory to describe a choice situation (the prospects themselves).

We performed our experiment in an Eastern-European country (Moldova) where we could use comparatively large stakes of up to a weekly salary. We analyze the data using Bayesian hierarchical modeling using a general theory of reference-dependent preferences that permits estimation of a variety of distinct types of reference point rules *ceteris paribus*, i.e. keeping all other behavioral parameters constant. Decision models are usually estimated through one of two approaches: the representative agent approach, which pools all data and assumes that all individuals are identical, or individual estimation, which assumes that all individuals are completely independent. Bayesian hierarchical modeling is a compromise between these two extremes. It estimates models at the individual level, but assumes that individuals share similarities and that their individual parameter values come from a population-level distribution. This permits less biased and more accurate parameter estimation and prevents inference from being dominated by outliers (Nilsson et al., 2011, Rouder and Lu, 2005).

The results indicate that most of our subjects take either the status quo or a security-based reference point (MaxMin - the maximum outcome they can be sure to obtain) as their reference point. Together these two reference points account for the behavior of over sixty percent of our subjects. Around twenty percent of our subjects use the prospect itself as a reference point, as suggested by Köszegi and Rabin (2006, 2007).

2. Theoretical background

A *prospect* is a *probability distribution* over a set of *outcomes* X , which is an interval of the reals. In our experiment outcomes are monetary payoffs. *Simple prospects* assign probability 1 to a finite set of outcomes. We denote these simple prospects as $(p_1, x_1; \dots; p_n, x_n)$, which means that they pay ϵx_j with probability $p_j, j = 1, \dots, n$. The decision maker has a weak preference relation \succsim over the set of prospects and, as usual, we denote strict preference by \succ , indifference by \sim , and the reversed preferences by \preceq and \prec . A real-valued evaluation function V *represents* \succsim if for all prospects $F, G, F \succsim G \Leftrightarrow V(F) \geq V(G)$.

Outcomes are defined as gains and losses relative to a reference point $r \in X$. An outcome x is a *gain* if $x > r$ and a *loss* if $x < r$. The set of outcomes recoded as gains and losses is $X_r = \{x \in \mathbb{R}: x + r \in X\}$.

2.1. Prospect theory

The main reference-dependent theory is prospect theory (Tversky and Kahneman, 1992). Under prospect theory, there exist *weighting functions* w^+ and w^- and a *gain-loss utility function* $U: X_r \rightarrow \mathbb{R}$ with $U(0) = 0$ and the evaluation function is given by

$$F \rightarrow PT_r(F) = \int_{x \geq r} U^+(x - r) dw^+(F) - \int_{x \leq r} U^-(x - r) dw^-(F) \quad (1)$$

where the integrals are Lebesgue integrals, and $U^+ = \max(U, 0)$ and $U^- = \max(-U, 0)$ are the positive and the negative parts of the utility function U .

U is an overall utility function and it includes loss aversion. The function w^+ is a probability weighting function for gains and w^- is a probability weighting function for losses. The probability weighting functions $w^s, s \in \{+, -\}$ map probabilities into $[0,1]$, they satisfy $w^s(0) = 0, w^s(1) = 1$, and they are *increasing*. When w^+ and w^- are linear, PT is equivalent to *expected utility*:

$$F \rightarrow EU_r(F) = \int_X U(x - r)dF . \quad (2)$$

Tversky and Kahneman hypothesized that U is S-shaped, concave for gains and convex for losses, and that the probability weighting functions w^s are inverse S-shaped, reflecting overweighting of small probabilities and underweighting of middle and large probabilities. Empirical evidence has generally confirmed these predictions (Wakker, 2010, pp.264-267).

2.2. Köszegi and Rabin's model

Tversky and Kahneman (1992) defined prospect theory for a riskless reference point r . Köszegi and Rabin (2006, 2007) added two elements to prospect theory. First, they made a distinction between the economic concept of outcome-based (or consumption) utility and the psychological concept of gain-loss utility and, second, they allowed the reference point to be random. Let R be the random reference point. In Köszegi and Rabin's model preferences over prospects F are represented by

$$F \rightarrow KR_R(F) = \int_X v(x)dF + \int_X \int_X U(v(x) - v(r))dFdR \quad (3)$$

where v represents consumption utility and U is the gain-loss utility function, which reflects the psychological part of utility. Consumption utility does not depend on the reference point and depends only on the absolute size of the payoffs.

Köszegi and Rabin (2007, p.1052) argue that “for modest-scale risk, such as \$100 or \$1000,[...] consumption utility can be taken to be approximately linear”. As the incentives in our experiment did not exceed \$1000 (in PPP) and the prospects in the different choice sets had approximately equal expected value, we concentrate on the gain-loss function U and take $v(x) = x$:

$$KR_R(F) = \int_X x dF + \int_X EU_r(F) dR . \quad (4)$$

The assumption of linear consumption utility is also common in empirical applications of Köszegi and Rabin's model (e.g. Abeler et al., 2011, Gill and Prowse, 2012, Heidhues and Köszegi, 2008, Eil and Lien, 2014).

While prospect theory is silent about the formation of the reference point, Köszegi and Rabin (2007) assume that it is based on the decision maker's rational expectations. They distinguish two specifications, one prospect-specific and one choice-specific. In a "*choice-acclimating personal equilibrium*" (CPE), the reference point is the prospect itself. This prospect-specific reference point implies that prospects are evaluated as:

$$KR(F) = \int_x x dF + \int_x EU_r(F) dF . \quad (5)$$

A choice-specific variant is Köszegi and Rabin's model with an "*unacclimating personal equilibrium*" (UPE), in which the reference point is the preferred prospect in the choice set. There is no probability weighting in Köszegi and Rabin's (2006, 2007) model. The main reason is that it is not clear how the rational expectations reference point is determined when there is probability weighting. Köszegi and Rabin (2006, 2007) do not explain this and abstract from probability weighting, even though they acknowledge its relevance (Köszegi and Rabin 2006, footnote 2, p. 1137).

2.3. *Disappointment models*

Köszegi and Rabin's (2006, 2007) model is close in spirit to the disappointment models of Bell (1985), Loomes and Sugden (1986), Gul (1991), and Delquié and Cillo (2006). Bell's model is equivalent to Eq. (3) with $v(r)$ replaced by the expected value of the prospect (although Bell remarks that this may be too restrictive and also presents a more general model), Loomes and Sugden's model (1986) is equivalent to Eq.(3) with

$v(r)$ replaced by the expected utility³ of the prospect, and Gul's (1991) model is equivalent to Eq.(3) with $v(r)$ replaced by the certainty equivalent of the prospect. Delquié and Cillo's (2006) model is identical to Köszegi and Rabin's (2007) CPE model (Eq. 5).

2.4 General reference-dependent specification

The purpose of this paper is to test how people's reference points are shaped when choosing between prospects. To explore this question statistically, it is necessary to use the same model specification across all reference point rules considered. In other words, in the estimation of the reference point, all other behavioral parameters must enter the model in the same way. To address this *ceteris paribus* principle we adopt the following general reference-dependent model:

$$F \rightarrow RD(F) = \int_X x dF + \int_X PT_r(F) dR. \quad (6)$$

Eq. (6) contains prospect theory (Eq. 1), Köszegi and Rabin's (2006, 2007) model (Eq. 5) and the disappointment models as special cases. In Eq. (6), probability weighting plays a role in the psychological part of the model, but it does not affect consumption utility. We believe this is reasonable as consumption utility reflects the "rational" part of utility and probability weighting is usually considered irrational. Adjusting the model to include probability weighting in the economic part is straightforward. Similarly, there is no weighting function when integrating the psychological part over the stochastic reference point R because the stochastic reference point is used to capture the decision maker's rational expectations.

³ With consumption utility v .

3. Reference point rules

A reference point rule specifies for each choice situation which reference point is used. Table 1 summarizes the reference point rules that we study in this paper. We distinguish reference point rules along two dimensions. First, whether they are *prospect-specific* (determine a reference point for each prospect separately), or *choice-specific* (determine a common reference point for all prospects within a choice set) and, second, whether they determine a random or a deterministic reference point.

	Prospect/Choice Specific	Random
Status Quo	Choice	No
MaxMin	Choice	No
MinMax	Choice	No
X at $MaxP$	Choice	No
Expected Value	Prospect	No
Prospect Itself	Prospect	Yes

Table 1. The reference point rules studied in this paper

The first candidate for the reference point is the *Status Quo*, which is often used in (experiments on) reference-dependent models. As subjects entered the lab having won nothing, we took the show-up fee as the status quo reference point and any extra money that subjects could win if one of their choices was played out for real was treated as a gain.

MaxMin, the second reference point rule, is based on Hershey and Schoemaker (1985). They found that when subjects were asked for the probability p that made them indifferent between outcome z for sure and a prospect $(p, x_1; 1 - p, x_2)$, they took z as their reference point and perceived $x_1 - z$ as a gain and $x_2 - z$ as a loss. Bleichrodt et al. (2001) and van Osch et al. (2004, 2006, 2008) found similar evidence for such a reference point formation strategy in medical decisions. Van Osch et al. (2006)

presented anecdotal evidence for this cognitive strategy. They asked their subjects to think aloud while making their choices. The most common reasoning in a choice between life duration z for sure and a prospect $(p, x_1; 1 - p, x_2)$ was: "I can gain $x - z$ years if the gamble goes well or lose $z - y$ if it doesn't."

The above reasoning implies that people are looking for security. In a comparison between two prospects, people look at the minimum outcomes of the two prospects and take the maximum of these as their reference point. This reference point is the amount they can obtain for sure. For example, in a comparison between $(0.50, 100; 0.50, 0)$ and $(0.25, 75; 0.75, 25)$, the minimum outcomes are 0 and 25 and because 25 exceeds 0 the rule predicts that subjects take 25 as their reference point.

MinMax is a bold counterpart of the cautious *MaxMin* rule. A *MinMax* decision maker looks at the maximal opportunities and takes the minimum of the maximum outcomes as his reference point. Hence, *MinMax* predicts that the decision maker takes 75 as his reference point when choosing between $(0.50, 100; 0.50, 0)$ and $(0.25, 75; 0.75, 25)$.

The *MaxMin* and the *MinMax* rules both look at the extreme outcomes. One reason is that these outcomes are salient. Another salient outcome is the payoff with the highest probability and the next rule, *XatMaxP*, defines this outcome as the reference point. The importance of salience is widely-documented in cognitive psychology (e.g. Kahneman 2011). Barber and Odean (2008) and Chetty et al. (2009) show the effect of salience on economic decisions. Bordalo et al. (2012) present a theory of salience in decision under risk.

The final two reference points that we considered are the *expected value* of the prospect, as in the disappointment models of Bell (1985) and Loomes and Sugden

(1986)⁴ and the *prospect itself* as in Köszegi and Rabin's (2007) CPE model and Delquié and Cillo's (2006) disappointment model. In contrast with the other reference points, these reference points are prospect-specific. The prospect itself is the only rule that specifies a random reference point. We do not consider the preferred prospect in a choice as in Köszegi and Rabin's (2007) UPE model, because the model in Eq. (6) is then defined recursively and could not be estimated.⁵

4. Experiment

Subjects

The subjects were 139 (49 females, age range 17-47, average age 22 years) students and employees from the Technical University of Moldova. They received a 50 Lei participation fee. In addition, to incentivize the experiment, subjects had a chance of one third to play out one of their choices for real.

The payoffs were substantial. The subjects who played out their choices for real earned 330 Lei on average, which was more than half the average weekly salary in Moldova at the time of the experiment. Two subjects won about 600 Lei, the average weekly salary.

Procedure

The experiment was computer-run in group sessions of 10 to 15 subjects. Subjects took 30 minutes on average to complete the experiment including instructions.

⁴ The equivalence with Loomes and Sugden (1986) follows because we assume $v(x) = x$.

⁵ Moreover, the model was typically not well-defined when the preferred prospect was the reference point. Let $G|F$ denote prospect G given that prospect F is the reference point. The CPE model requires that either $F|F \succcurlyeq G|G \succcurlyeq G|F$ or $G|G \succcurlyeq F|F$. It is straightforward to construct choice pairs that, under standard behavioral parameters assumptions, exhibit the paradoxical pattern $F|F > G|G$ and $G|F > F|F$.

Subjects made 70 choices in total. Each choice involved two options, Option 1 and Option 2. Each option had between one and four possible outcomes. We randomized the order of the choices and we also randomized whether a prospect was presented as Option 1 or as Option 2.

Because we are interested in the Maxmin and the Minmax rules, we constructed choices with different maxima of the minima and with different minima of the maxima. Choices were created by an optimal design procedure that minimized the joint correlation between choices. The advantage of using minimally correlated choices is that they give more precise and more robust estimates of the behavioral parameters.

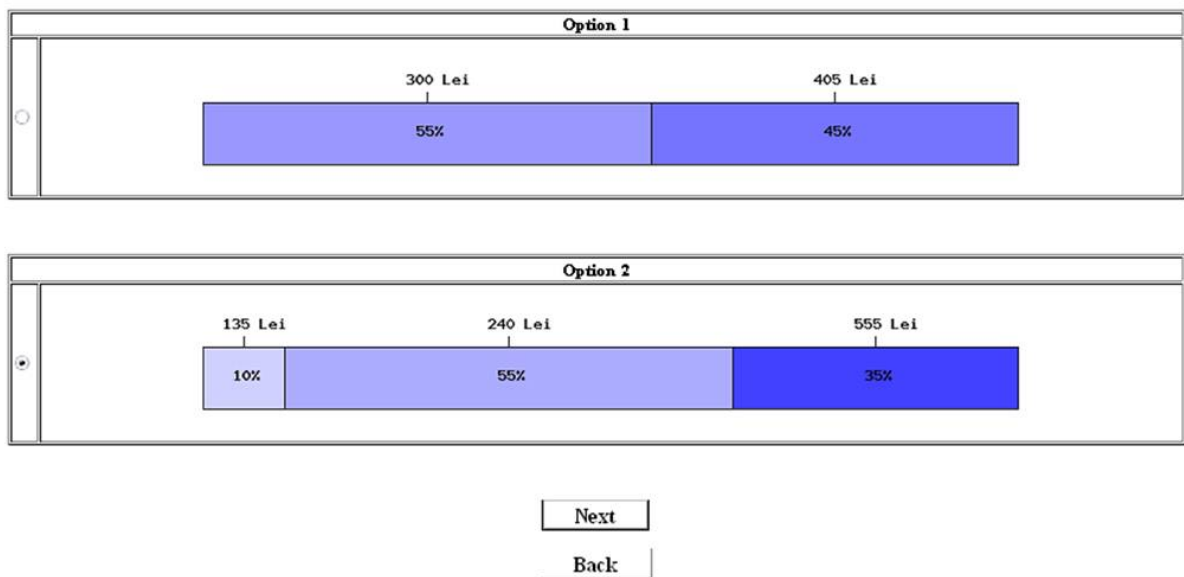


Figure 1. Presentation of the choices in the experiment

Figure 1 shows the presentation of the choices. Prospects were presented as horizontal bars with as many parts as there are different payoffs. The size of each part corresponded with the probability of the payoff and the intensity of the color (blue) was proportional to the size of the payoff. The payoffs were presented in increasing order.

Subjects were asked to click on a bullet to indicate their preferred option (Figure 1 illustrates a choice for Option 2).

5. Bayesian hierarchical modeling

We analyzed the data using Bayesian hierarchical modeling. Hierarchical modeling is an appealing compromise between assuming a representative agent (thus ignoring individual heterogeneity) and treating all individuals as independent. Hierarchical models assume that the individuals share similarities and that their individual parameter values come from a common (population-level) distribution. Hence, the parameter estimates for one individual benefit from the information that is obtained from all other participants. Hierarchical models simultaneously account for similarities and differences between individuals: individual parameters are estimated separately, but they are constrained by the higher-order common distribution (which is also estimated).

The common approach in behavioral decision making is to assume that all subjects are independent and to estimate the individual parameters by only using the choice data of that individual. Hierarchical modeling has two advantages over this approach. First, the data from the other respondents can be used in the estimation of the individual parameters, which improves the precision with which these parameters can be estimated. In Bayesian statistics this phenomenon is known as *collective inference*. A second advantage is that outliers have less impact on the results. Each individual parameter is shrunk towards the group mean, an effect that is stronger for individuals with noisier behavior, thus making the overall estimation more robust. This is particularly true for parameters that are estimated with lower precision. An example is the loss aversion coefficient in prospect theory, for which the standard deviation of the

parameter estimates is usually high. Nilsson et al. (2011) illustrate that Bayesian hierarchical modeling leads to more accurate and more efficient estimates of loss aversion than the commonly-used maximum likelihood estimation.

Figure 2 shows a schematic representation of our statistical model. The model consists of two main parts: the specification of the core behavioral parameters and the specification of the reference point rule.

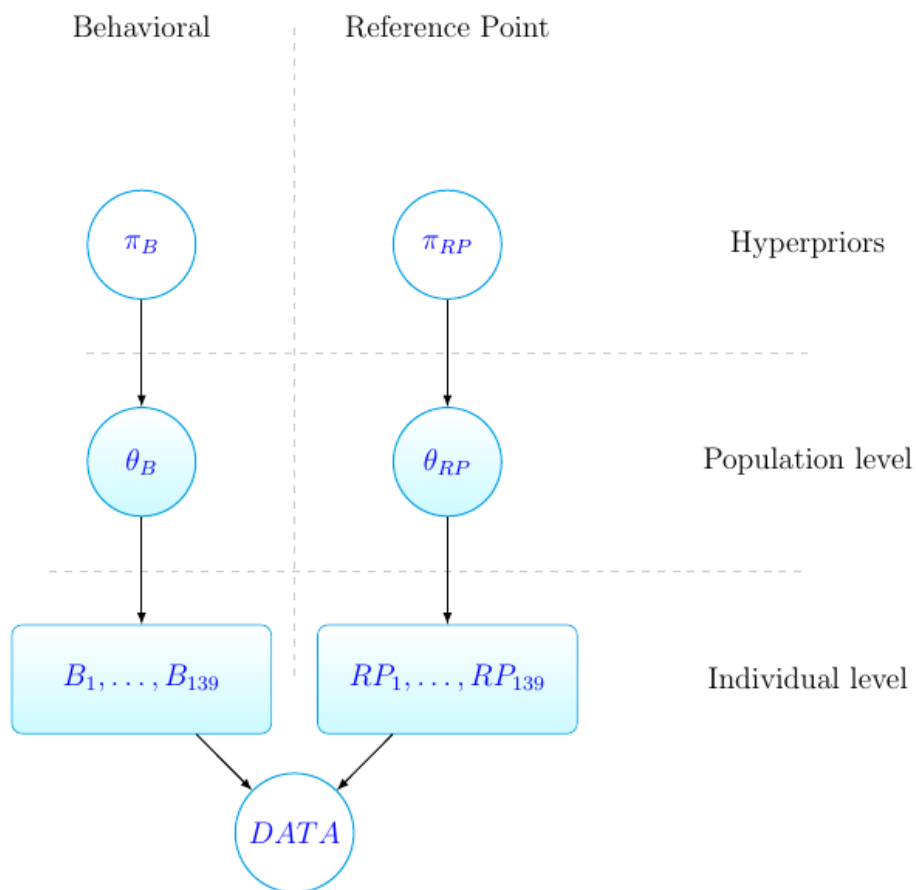


Figure 2. Graphical representation of our model. Empty nodes are known or predefined quantities, filled nodes are unknown latent parameters of our model.

We adopt the following mnemonic conventions. For individual $i \in \{1, \dots, 139\}$ the vector of core behavioral parameters is denoted B_i and the reference point rule that he adopts is denoted RP_i . We assume that a subject uses the same reference point rule in all questions, where the reference point rule is one of the candidates listed in Table 1. The distribution of the core behavioral parameters in the population is parameterized by an unknown vector θ_B , and the distribution of reference point rules in the population is parameterized by θ_{RP} . The known hyper-priors, which are necessary for the complete specification of the Bayesian model, are denoted by π_B and π_{RP} respectively. The vector of the observed choices (data) of the individual i is denoted by $D_i = (D_{i1}, \dots, D_{i70})$. We will now describe our estimation procedure in detail.

5.1. Specification of the behavioral parameters

We assume that the utility function U in Eq. (6) is a power function:

$$U(x) = \begin{cases} (x - r)^\alpha & \text{if } x \geq r \\ -\lambda(r - x)^\alpha & \text{if } x < r \end{cases} \quad (7)$$

In Eq. (7) the α -parameter reflects the curvature of utility and the λ -parameter indicates loss aversion. We assumed the same curvature for gains and losses, because it is hard to estimate loss aversion when utility curvature for gains and for losses can both vary freely (Nilsson et al., 2011).

For probability weighting, we assumed Prelec's (1998) one-parameter specification:

$$w^s(p) = \exp(-(-\ln p)^\nu), s \in \{+, -\}. \quad (8)$$

We assume the same probability weighting for gains and losses. Empirical studies usually find that the differences in probability weighting between gains and losses are relatively minor (Tversky and Kahneman, 1992, Abdellaoui, 2000).

To account for the probabilistic nature of people's choices we used Luce's (1959) logistic choice rule. Let $RD(F)$ and $RD(G)$ denote the respective values of prospects F and G according to our general reference-dependent model, Eq. (6). Luce's rule says that the probability $P(F, G)$ of choosing prospect F over prospect G equals

$$P(F, G) = \frac{1}{1 + e^{\xi[RD(G) - RD(F)]}} . \quad (9)$$

In Eq. (9), $\xi > 0$ is a precision parameter that measures the extent to which the decision maker's choices are determined by the differences in value between the prospects. In other words, the ξ -parameter signals the quality of the decision. Larger values of ξ imply that choice is driven more by the value difference between prospects F and G . If $\xi = 0$, choice is random and if ξ goes to infinity choice essentially becomes deterministic. In his comprehensive exploration of prospect theory specifications, Stott (2006) concluded that power utility, the Prelec one-parameter probability weighting function, and Luce's choice rule gave the best fit to his data and we, therefore, selected these specifications. As robustness checks, we also ran our analysis with exponential utility, Prelec's (1998) two-parameter specification of the weighting function, and an alternative, more flexible weighting function.

Each of the 139 subjects in the experiment had his own parameter vector $B_i = (\alpha_i, \gamma_i, \lambda_i, \xi_i)$. Hierarchical models assume that although these parameters vary across individuals, they are drawn from a common population-level parent distribution (Rouder and Lu, 2005). The population-level distribution is not fully specified but has free parameters that are estimated from the data. These parameters also follow a distribution, but with a known shape. The specification of this distribution is the final layer in the hierarchical specification; it is commonly referred to as a hyper-prior.

We assumed that each parameter in B_i comes from a lognormal distribution: $\alpha_i \sim \text{logN}(\mu_\alpha, \sigma_\alpha^2)$, $\lambda_i \sim \text{logN}(\mu_\lambda, \sigma_\lambda^2)$, $\gamma_i \sim \text{logN}(\mu_\gamma, \sigma_\gamma^2)$, and $\xi_i \sim \text{logN}(\mu_\xi, \sigma_\xi^2)$. Thus, the complete vector of unknown parameters at the population-level is $\theta_G = (\mu_\alpha, \mu_\lambda, \mu_\gamma, \mu_\xi, \sigma_\alpha^2, \sigma_\lambda^2, \sigma_\gamma^2, \sigma_\xi^2)$. For the hyper-priors, $\pi_* = (\mu_*, \sigma_*^2)$, $*$ $\in \{\alpha, \lambda, \gamma, \xi\}$ of the parent distributions we made the usual assumption that the μ_* follow a lognormal distribution and that the σ_*^2 follow an inverse Gamma distribution. We centered the hyper-priors around linearity (expected value) and chose the variances such that the hyper-priors were diffuse and would have a negligible impact on the posterior estimation.

Denote the set of hyper-priors for the core behavioral parameters by π_B . Then the joint probability distribution of the behavioral parameters $\mathbf{B} = (B_1, \dots, B_{139})$ and θ_B equals

$$P(\mathbf{B}, \theta_B | \pi_B) = \left(\prod_{i=1}^{139} P(B_i | \theta_B) \right) P(\theta_B | \pi_B). \quad (10)$$

Given reference point rule RP_i , the likelihood of subject i 's responses equals

$$P(D_i | B_i, RP_i) = \prod_{q=1}^{70} P(D_{i,q} | G_i, RP_i). \quad (11)$$

The probability of each choice $D_{i,q}$ is computed using Luce's rule, Eq.(9). From Eqs. (10) and (11), it follows that the joint probability distribution of all the unknown behavioral parameters \mathbf{B} and θ_B and all the observed choices $\mathbf{D} = (D_1, \dots, D_{139})$ is

$$P(\mathbf{D}, \mathbf{B}, \theta_B | \mathbf{RP}, \pi_B) = \left(\prod_{i=1}^{139} \prod_{q=1}^{70} P(D_{i,q} | B_i, RP_i) \right) \left(\prod_{i=1}^{139} P(B_i | \theta_B) \right) P(\theta_B | \pi_B). \quad (12)$$

In Eq.(12), $\mathbf{RP} = (RP_1, \dots, RP_{139})$ is the vector of individual reference point rules.

4.2. Specification of the reference point rule

We assume that subjects use one of the six reference point rules specified in Table 1. To answer the question which of these six rules they used, we have to estimate the reference point rules from the data. In other words, we are interested in the posterior probability that an individual uses one of the six reference points given the data: $P(RP_i|\mathbf{D})$. To compute this probability, we must supplement Eq. (12) with the specification of the prior for the reference point rule.

RP_i is a six-dimensional categorical variable and it is conventional to use the Dirichlet distribution: $\theta_{RP} \sim \text{Dirichlet}(\pi_{RP})$, where θ_{RP} is a probability vector in a six-dimensional simplex and π_{RP} is a diffuse hyper-prior parameter for the Dirichlet distribution. Then the joint probability density of \mathbf{RP} and θ_{RP} becomes:

$$P(\mathbf{RP}, \theta_{RP} | \pi_{RP}) = (\prod_{i=1}^{139} P(RP_i | \theta_{RP})) P(\theta_{RP} | \pi_{RP}). \quad (13)$$

Substituting Eq.(13) into Eq.(12) gives the complete specification of our statistical model:

$$P(\mathbf{D}, \mathbf{B}, \theta_B, \mathbf{RP}, \theta_{RP} | \pi_B, \pi_{RP}) = (\prod_{i=1}^{139} \prod_{q=1}^{70} P(D_{i,q} | B_i, RP_i)) (\prod_{i=1}^{139} P(B_i | \theta_B)) (\prod_{i=1}^{139} P(RP_i | \theta_{RP})) P(\theta_B | \pi_B) P(\theta_{RP} | \pi_{RP}). \quad (14)$$

4.3. Estimation

To compute the marginal posterior distributions $P(B_i | \mathbf{D}, \pi_B, \pi_{RP})$, $P(RP_i | \mathbf{B}, \pi_B, \pi_{RP})$, $P(\theta_B | \mathbf{D}, \pi_B, \pi_{RP})$, and $P(\theta_{RP} | \mathbf{D}, \pi_B, \pi_{RP})$, we must integrate the various conditional distributions. Direct computation of these marginals is intractable and we therefore used Markov Chain Monte Carlo (MCMC) sampling (Gelfand and Smith, 1990) with blocked Gibbs sampling.⁶ We first used 10,000 burn-in iterations with adaptive MCMC

⁶ For the behavioral parameters B_1, \dots, B_{139} we used Metropolis-Hasting MCMC with symmetric normal proposal on the log-scale, for the block RP_1, \dots, RP_{139} we used Metropolis-Hasting MCMC with

and then 20,000 standard MCMC burn-in iterations. The results are based on the subsequent 50,000 iterations.

5. Results

5.1. Consistency

To test for consistency, five choices were asked twice. In 68.7% of these repeated choices, subjects twice made the same choice. Previous evidence has indicated that reversal rates up to one third are common in experiments (Stott, 2006). Moreover, our choices were complex, involving more than two outcomes and with expected values that were close.

5.2. Reference points

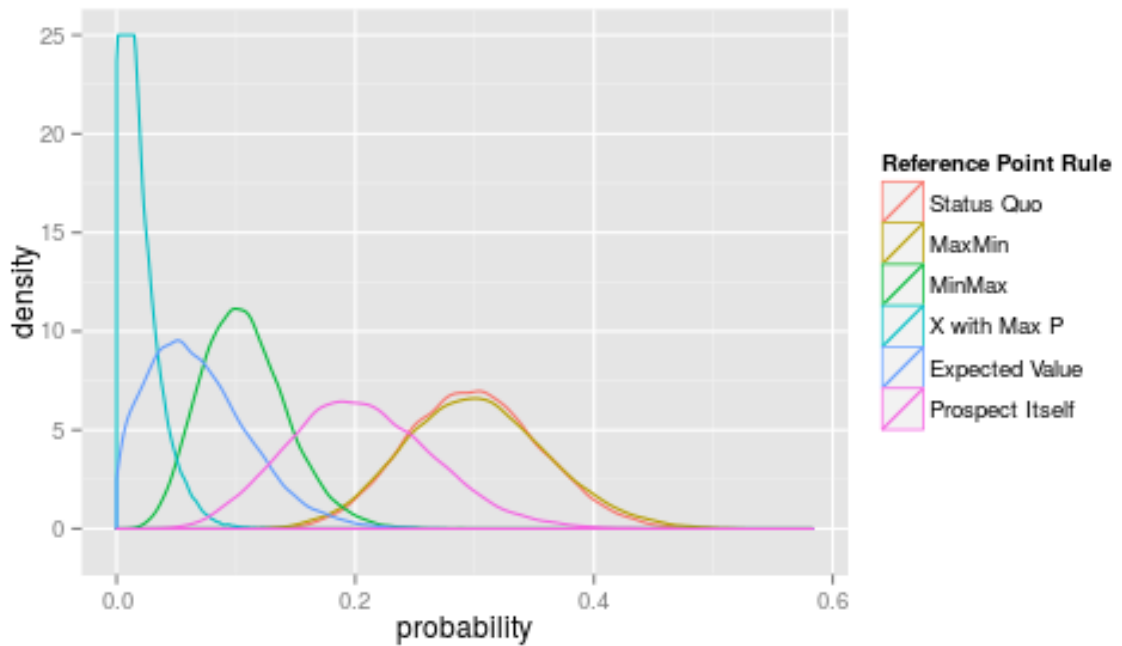


Figure 3. Marginal posterior distributions of each reference point rule

uniform proposal, and the group-level blocks θ_G and θ_{RP} are sampled directly from the conjugate Gamma-Normal and Dirichlet-Categorical distributions, respectively.

We first report our estimates of θ_{RP} , the probabilities of a randomly chosen subject being best described by each reference point rule. Figure 3 shows the marginal posterior distributions of θ_{RP} in the population, for each RP rule. Table 2 reports the medians and standard deviations of these distributions.⁷

	Median	SD
StatusQuo	0.30	0.06
MaxMin	0.30	0.06
MinMax	0.10	0.04
X with Max P	0.01	0.02
ExpectedValue	0.06	0.04
Prospect Itself	0.20	0.06

Table 2. Point estimates of RP mixture in the population

The reference points with the highest probabilities were the Status Quo and the MaxMin rule according to which subjects use the minimum outcome they could be sure of as their reference point. According to our median estimates, each of these two rules was used by 30% of the subjects. The prospect itself (the rule suggested by Köszegi and Rabin (2006, 2007) and Delquié and Cillo (2006)) was used by 20% of the subjects. The other three rules were used rarely.

At the individual subject level, we can also assess the likelihood that a subject uses a specific reference point by looking at the posterior distributions. Figure 4 shows, for example, the posterior distributions for subjects 17, 50, and 100. Subject 17 has about 60% probability to use the prospect itself as his reference point and 25% probability to use the minimum of the maximums. Subject 50 clearly uses the maximum amount he can be sure of as his reference point, and subject 100 clearly uses the status quo as his reference point.

⁷ Note that the medians need not add to 100%.

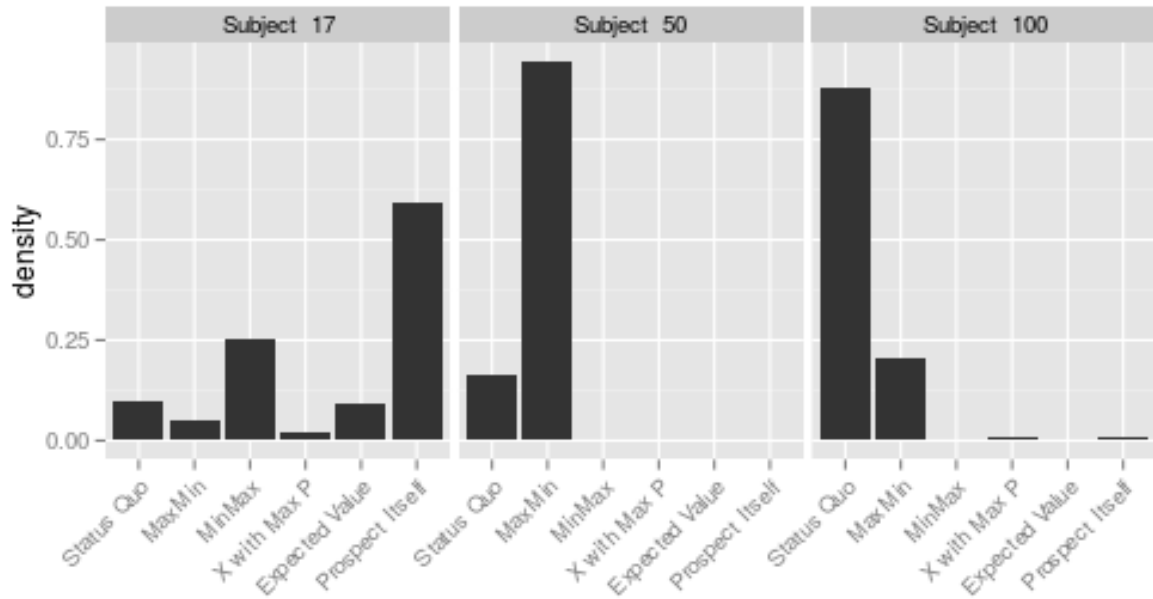


Figure 4. Posterior distributions for subjects 17, 50, and 100.

We say that a subject is classified *sharply* if, at the individual level, any of the six reference points has a posterior classification probability of at least 50%. For example, subjects 17, 50, and 100 were all classified sharply. Out of 139 subjects, 107 could be classified sharply. Figure 5 shows the distribution of responses over the six reference point rules for the sharply classified subjects. The dominance of the Status Quo and the MaxMin rule increased further and around 70% of the sharply classified subjects used one of these two reference points. The proportions of the other rules decreased slightly and the little support for the X at maxP rule and for the expected value rule that we observed before all but disappeared.

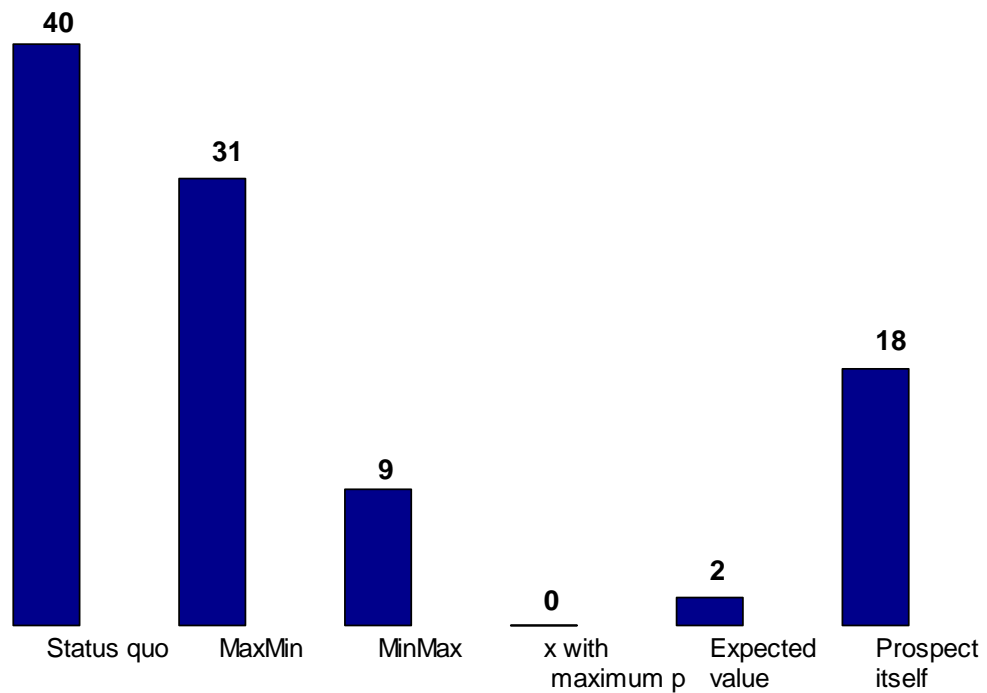


Figure 5. Proportion of sharply classified respondents satisfying a particular reference point rule (percent)

5.3. Behavioral parameters

Figure 6 shows the gain-loss utility function based on the estimated behavioral population level parameters (θ_B) in the psychological (PT) part of Eq. (4). The utility function has the usual S-shape: concave for gains and convex for losses. We found more utility curvature than most previous estimations of gain-loss utility (Fox and Poldrack, 2014). Our estimated utility function is close to the functions estimated by Wu and Gonzalez (1996), Gonzalez and Wu (1999), and Toubia et al. (2013). The loss aversion coefficient equaled 2.34, which is in line with other findings in the literature. Appendix A shows the posterior densities of all behavioral population level parameters.

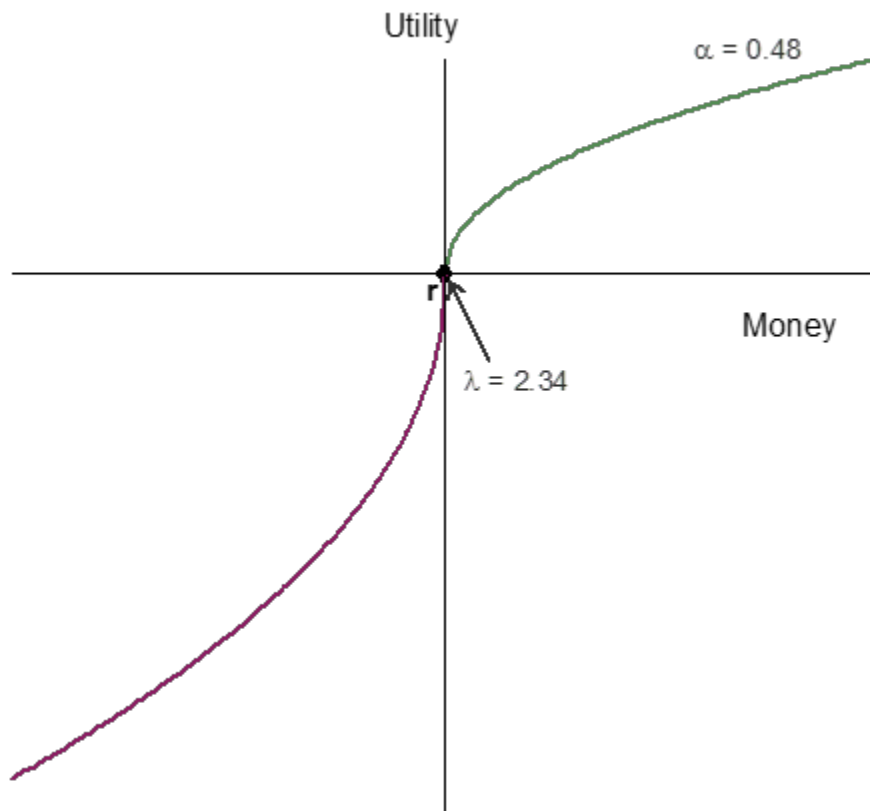


Figure 6. The utility function based on the estimated group parameters

Figure 7 shows the estimated probability weighting function in the population. The function has the commonly observed inverse S-shape, which reflects overweighting of small probabilities and underweighting of intermediate and large probabilities. Our estimated probability weighting function is close to the functions that were estimated by Gonzalez and Wu (1999), Bleichrodt and Pinto (2000) and Toubia et al. (2013).

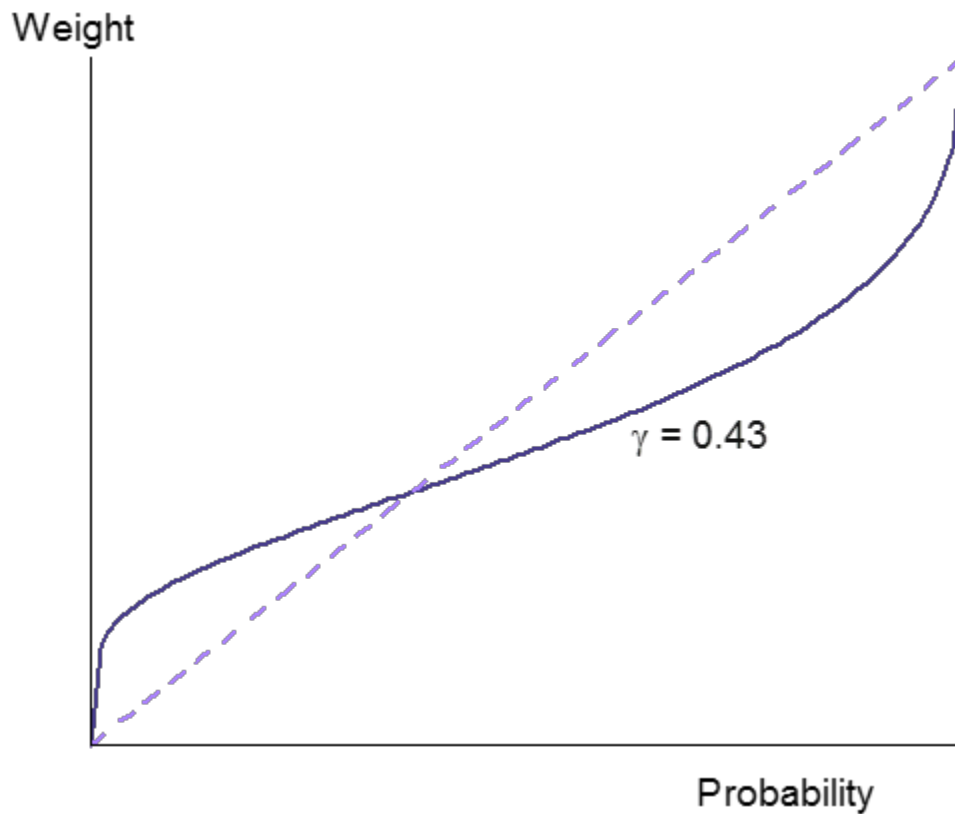


Figure 7. The probability weighting function based on the estimated group parameters

As mentioned earlier, an important advantage of Bayesian hierarchical modeling is that it permits expressing uncertainty in the individual parameter estimates by means of the posterior densities. To illustrate, Figure 8 shows the posterior densities of subject 17. As the graph shows, subject 17's parameter estimates varied considerably, although it is safe to say that he had concave utility and inverse S-shaped probability weighting.

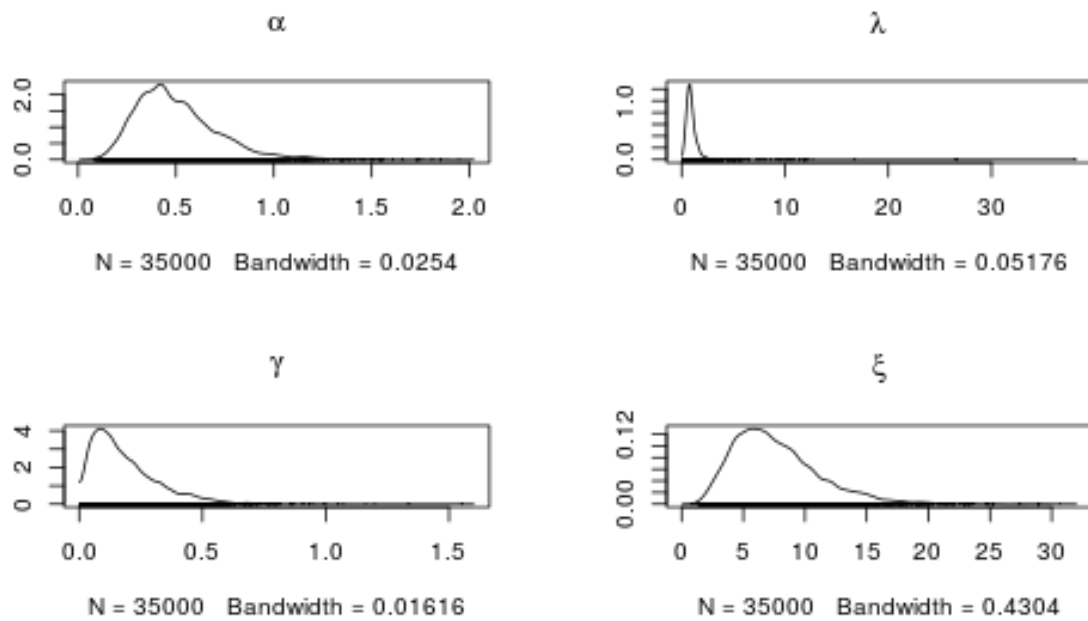


Figure 8. Posterior densities of the behavioral parameters for subject 17.

Table 3 shows the quantiles of the posterior point estimates of all 139 subjects. The table shows that utility curvature and, to a lesser extent, probability weighting were rather stable. Loss aversion varied, however, much more although for more than 75% of the subjects, the estimate was in line with loss aversion.

	2.5%	25%	50%	75%	97.5%
α	.31	.40	.44	.50	.60
γ	.09	.14	.24	.44	1.66
λ	.36	1.19	1.59	2.25	4.63
ξ	6.11	8.26	10.89	14.41	25.76

Table 3. Quantiles of the point estimates of the behavioral parameters for all 139 subjects

Table 4 shows the median behavioral parameters for the sharply classified subjects in each group. A priori, it seemed plausible that subjects who used different reference points might also have different behavioral parameters and the table shows evidence for this. While utility curvature and probability weighting were rather stable across the

groups, the loss aversion coefficients varied from 0.50 in the MinMax group to 2.44 in the Expectation group. The loss aversion coefficient of 0.50 in the MinMax group has the interesting interpretation that these optimistic subjects weight gains twice as heavy as losses and they exhibit what might be seen as the reflection of the preferences of the cautious MaxMin subjects.

	α	γ	λ	ξ
Status quo	.42	.28	1.51 ⁸	11.75
MaxMin	.46	.24	2.24	10.30
MinMax	.40	.15	.50	14.34
Expectation	.36	.25	2.44	6.14
Prospect	.45	.16	2.23	10.89

Table 4: Median individual level parameters for the sharply classified subjects in each group.

Table 4 also shows that subjects who used the status quo as their reference point were typically no expected utility maximizers as there was substantial probability weighting in the status quo group. Table 5 shows the subdivision of the subjects who used the status quo as their reference point based on the 95% Bayesian credible intervals of their estimated utility curvature and probability weighting parameters. Twelve subjects (those with $\gamma = 1$) behaved according to expected utility, three of which (those with $\alpha = 1$ and $\gamma = 1$) were expected value maximizers. Consequently, less than 10% of our subjects were expected utility maximizers.

⁸ The reason λ is not equal to 1 for subjects who were sharply classified as using the status quo rule is that a subject's behavioral parameters stayed the same for all reference point rules. Consequently, even when a subject was (sharply) classified as a status quo type, there was still a non-negligible probability that he used any of the other reference point rules.

		Probability weighting			Total
		$\gamma < 1$	$\gamma = 1$	$\gamma > 1$	
Utility	$\alpha < 1$	28	9	0	37
	$\alpha = 1$	3	3	0	9
	$\alpha > 1$	0	0	0	0
	Total	31	12	0	43

Table 5. Behavioral parameters of the subjects using status quo as their reference points.

5.4. Robustness

Throughout our main analysis, we used the utility and weighting functions given in Eq. (7) and Eq. (8). We performed three robustness checks, replacing power utility by exponential utility, Prelec’s (1998) one-parameter weighting function by his two-parameter function, and finally using another more flexible weighting function.⁹ In all three cases, the status quo and Maxmin remained by far the most dominant reference points, always capturing the behavior of more than 60% of the subjects. Expectations-based reference point performed worse. Detailed results are reported in the online appendix.

Up to now, we assumed Eq. (4) for all reference point rules, allowing us to keep all behavioral parameters constant when comparing reference point. To further test the robustness of our findings we also tried several other specifications, which are summarized in Table 6. Model 1 corresponds to the results reported in Sections 5.2 and 5.3. The two variables we varied in these robustness checks were the inclusion of consumption utility and probability weighting. While models with prospect-specific reference points need consumption utility to exclude implausible choice behavior (e.g. the choice of a clearly dominated prospect), models with a choice-specific reference point do not. Prospect theory, for example, does not include consumption utility.

⁹ We used the incomplete Beta function because it allows for a wide variety of shapes. To our knowledge, this function has only been used to model probability weighting by Wilcox (2012).

Consequently, we estimated the models with a choice-specific reference point both with and without consumption utility.

Model	Choice-specific reference point		Prospect-specific reference point	
	Consumption utility	Probability weighting	Consumption utility	Probability weighting
1	Yes	Yes	Yes	Yes
2	No	Yes	Yes	Yes
3	Yes	Yes	Yes	No
4	No	Yes	Yes	No
5	Yes	No	Yes	No
6	No	No	Yes	No

Table 6: Estimated models

In Eq. (4) with a prospect-specific reference point, we assumed that subjects apply probability weighting in the overall evaluation, but, following the literature on stochastic reference points (Kőszegi and Rabin, 2006, Kőszegi and Rabin, 2007, Delquié and Cillo, 2006, Sugden, 2003, Schmidt et al., 2008), we abstracted from probability weighting in the determination of the stochastic reference point. It is particularly difficult to see how rank-dependence can be defined in a model in which probabilities are multiplied. This raises similar questions as those posed in the literature on dynamic choice under non-expected utility and which have not been answered satisfactorily yet. However, to have probability weighting in the overall evaluation but not in the determination of the reference point may be arbitrary and we, therefore also estimated the models without probability weighting. We performed two sets of estimations: one in which the models with a choice-specific reference point included probability weighting, but the models with a prospect-specific reference point did not (models 3 and 4) and one in which all model had no probability weighting (models 5 and 6).

The results of the robustness checks were as follows. First, our main conclusion that the status quo and MaxMin were the dominant reference points remained valid. The behavior of 60% to 75% of the subjects was best described by a model with one of these two reference points. Second, it is important to take probability weighting into account. Excluding probability weighting from the models with a prospect-specific reference point (model 3) increased the share of MaxMin reference point to 44% (52% if we only include sharply classified subjects) and the share of the prospect itself as a reference point decreased to 10% (8% if we only include sharply classified subjects). The other shares changed only little. Prospect-specific models like those of Loomes and Sugden (1986) and Köszegi and Rabin's (2006, 2007) benefit from the inclusion of probability weighting. Finally, ignoring probability weighting altogether, as in models 5 and 6, led to unstable estimation results.

The behavioral parameters were comparable across all models that we estimated. The power utility coefficient was approximately 0.50 in all models, the probability weighting parameter varied between 0.40 and 0.60 (except, of course, when no probability weighting was assumed), and the loss aversion coefficient varied between 2 and 2.50. Excluding consumption utility from models with a choice-specific reference point (models 2 and 4) led to a substantial increase in the precision parameter ξ . Full results of the robustness analysis are available in the online appendix.

6. Discussion

Empirical evidence shows that reference points have a substantial effect on people's preferences. However, there is little insight into how reference points are formed. Our results indicate that for most subjects the reference point was determined by the choice situation and they took either the status quo or a security-level reference point (the

payoff the decision maker can be sure to obtain) as their reference point. There is much less support for expectations-based reference points as in the disappointment models and Köszegi and Rabin's (2006, 2007) CPE model.

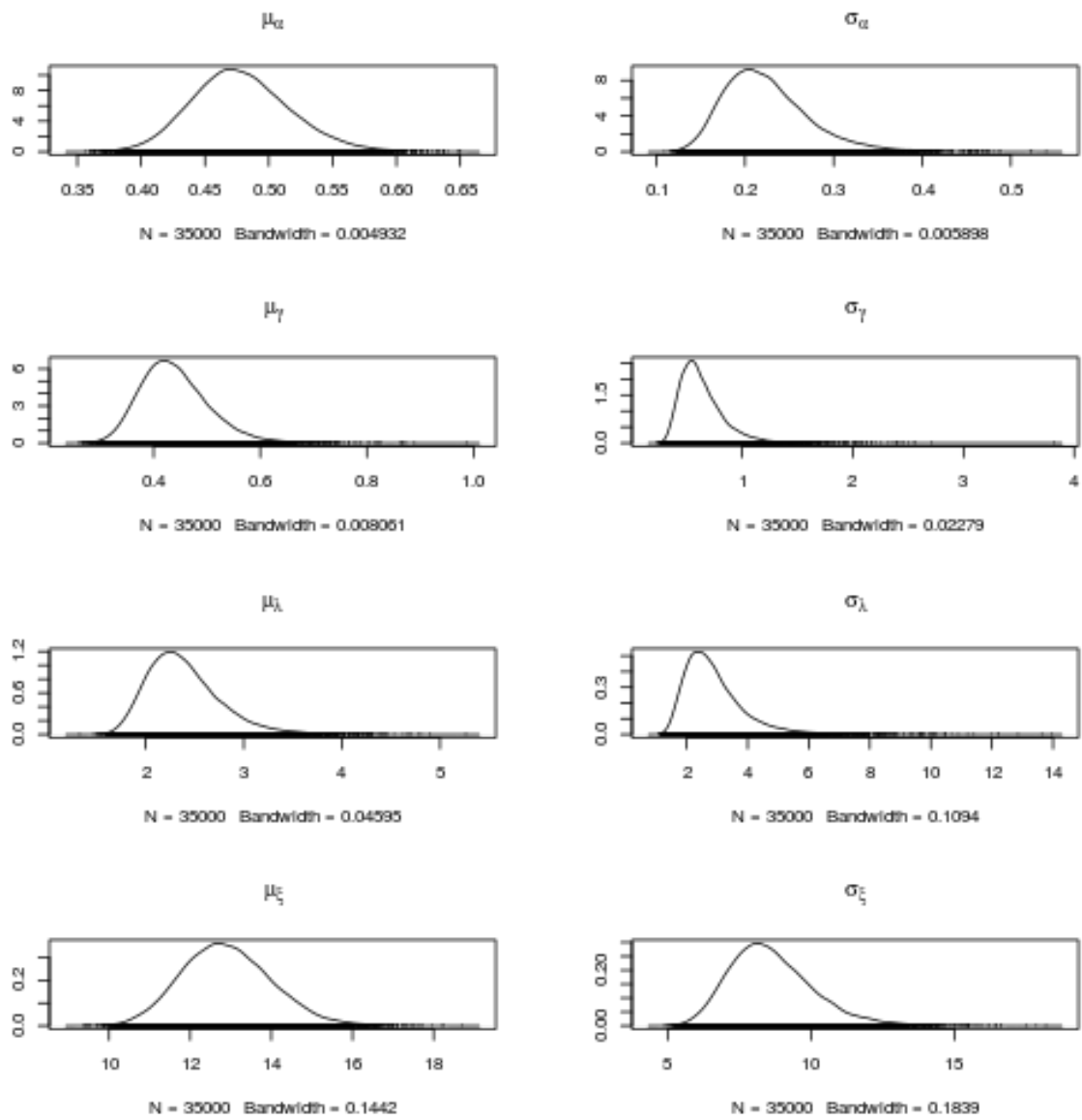
Experiments in decision under risk often assume that subjects take the status quo (0) as their reference point. Our data show that this assumption is justified for 30-40% of the subjects, but that a majority uses a different reference point. Our data also suggest how a researcher can increase the likelihood that subjects use 0 (or the show-up fee) as their reference point. For example, in choosing between mixed prospects, researchers could include a prospect with 0 as its minimum outcome in subjects' choice set. This ensures that MaxMin subjects will also use 0 as their reference point and, as our results suggest, that a substantial majority of the subjects will use 0 as their reference point. Our results also assess the validity of earlier empirical studies that took the status quo as the reference point.

We tested the reference point rules in an experiment with large incentives (subjects could win up to a weekly salary) and used a Bayesian hierarchical approach to analyze the data. Economic researchers usually estimate models either by pooling all data or by individual estimation. Both approaches have their limitations. Pooling ignores heterogeneity between individual decision makers and may result in estimates that are not representative of any individual in the sample. In individual estimation subjects contribute only few data and this may lead to unreliable parameter estimates. Bayesian analysis strikes a nice balance between pooling and individual estimation and it leads to more precise parameter estimates. A potential limitation of Bayesian analysis is that the selected priors can affect the estimations, but in our analysis the choice of priors had negligible impact on the estimates.

To make inferences about the different reference point rules, we used a general theoretical framework which allowed to vary the reference point rule, while keeping all other behavioral variables constant across the rules (*ceteris paribus* principle). This approach is cleaner and better interpretable than standard mixture modeling, which estimates the weights of different models globally. Using a Bayesian model has the additional advantage that we could obtain the parameter estimates both for the distribution of reference point rules in the population and for each subject separately.

We did not test all reference points that have been proposed in the literature. For example, we did not test explicitly for subjects' goals. On the other hand, subjects may have had few goals for the current experiment and it is also possible that their goals were equal to one of the reference points that we used (e.g. expected value or the security level). A similar remark applies to aspiration levels. Testing for goals and aspiration levels seems difficult within the revealed preference paradigm that we used and may require other data inputs than choice data alone.

Appendix A: The posterior densities of the behavioral parameters



References

- Abdellaoui, M. (2000), 'Parameter-free elicitation of utility and probability weighting functions', *Management Science*, vol. 46, pp. 1497-1512.
- Abeler, J., Falk, A., Goette, L., Huffman, D. (2011), 'Reference points and effort provision', *American Economic Review*, vol. 101, pp. 470-492.
- Allen, E.J., Dechow, P.M., Pope, D.G., Wu, G. (2013), 'Reference-dependent preferences: Evidence from marathon runners' *Working Paper*.
- Barber, B.M., Odean, T. (2008), 'All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors', *Review of Financial Studies*, vol. 21, pp. 785-818.
- Barberis, N.C. (2013), 'Thirty years of prospect theory in economics: A review and assessment', *Journal of Economic Perspectives*, vol. 27, pp. 173-196.
- Bartling, B., Brandes, L., Schunk, D. (2015), 'Expectations as Reference Points: Field Evidence from Professional Soccer', *Management Science*, vol. 61, pp. 2646-2661.
- Baucells, M., Weber, M., Welfens, F. (2011), 'Reference-point formation and updating', *Management Science*, vol. 57, pp. 506-519.
- Bell, D.E. (1985), 'Disappointment in decision making under uncertainty', *Operations Research*, vol. 33, pp. 1-27.
- Benartzi, S., Previtro, A., Thaler, R.H. (2011), 'Annuity puzzles', *Journal of Economic Perspectives*, vol. 25, pp. 143-164.

- Benartzi, S., Thaler, R.H. (1995), 'Myopic loss aversion and the equity premium puzzle', *Quarterly Journal of Economics*, vol. 110, pp. 73-92.
- Bleichrodt, H., Pinto, J.L. (2000), 'A parameter-free elicitation of the probability weighting function in medical decision analysis', *Management Science*, vol. 46, pp. 1485-1496.
- Bleichrodt, H., Pinto, J.L., Wakker, P.P. (2001), 'Making descriptive use of prospect theory to improve the prescriptive use of expected utility', *Management Science*, vol. 47, pp. 1498-1514.
- Bordalo, P., Gennaioli, N., Shleifer, A. (2012), 'Salience theory of choice under risk', *Quarterly Journal of Economics*, vol. 127, pp. 1243-1285.
- Card, D., Dahl, G.B. (2011), 'Family violence and football: The effect of unexpected emotional cues on violent behavior', *Quarterly Journal of Economics*, vol. 126, pp. 103-143.
- Chetty, R., Looney, A., Kroft, K. (2009), 'Salience and taxation: Theory and evidence', *American Economic Review*, vol. 99, pp. 1145-1177.
- Crawford, V.P., Meng, J. (2011), 'New york city cab drivers' labor supply revisited: Reference-dependent preferences with rational expectations targets for hours and income', *American Economic Review*, vol. 101, pp. 1912-1932.
- Delquié, P., Cillo, A. (2006), 'Disappointment without prior expectation: A unifying perspective on decision under risk', *Journal of Risk and Uncertainty*, vol. 33, pp. 197-215.

- Diecidue, E., Van de Ven, J. (2008), 'Aspiration level, probability of success and failure, and expected utility', *International Economic Review*, vol. 49, pp. 683-700.
- Eil, D., Lien, J.W. (2014), 'Staying ahead and getting even: Risk attitudes of experienced poker players', *Games and Economic Behavior*, vol. 87, pp. 50-69.
- Fox, C.R., Poldrack, R.A. (2014). 'Prospect theory and the brain', in (P. Glimcher, E. Fehr, eds.), *Handbook of Neuroeconomics (2nd Ed.)*, pp. 533-567, New York: Elsevier.
- Gelfand, A.E., Smith, A.F. (1990), 'Sampling-based approaches to calculating marginal densities', *Journal of the American Statistical Association*, vol. 85, pp. 398-409.
- Genesove, D., Mayer, C.J. (2001), 'Loss aversion and seller behavior: Evidence from the housing market', *Quarterly Journal of Economics*, vol. 116, pp. 1233-1260.
- Gill, D., Prowse, V. (2012), 'A structural analysis of disappointment aversion in a real effort competition', *American Economic Review*, vol. 102, pp. 469-503.
- Gonzalez, R., Wu, G. (1999), 'On the form of the probability weighting function', *Cognitive Psychology*, vol. 38, pp. 129-166.
- Gul, F. (1991), 'A theory of disappointment aversion', *Econometrica*, vol. 59, pp. 667-686.
- Heath, C., Larrick, R.P., Wu, G. (1999), 'Goals as reference points', *Cognitive Psychology*, vol. 38, pp. 79-109.
- Heidhues, P., Köszegi, B. (2008), 'Competition and price variation when consumers are loss averse', *American Economic Review*, pp. 1245-1268.

- Hershey, J.C., Schoemaker, P.J.H. (1985), 'Probability versus certainty equivalence methods in utility measurement: Are they equivalent?', *Management Science*, vol. 31, pp. 1213-1231.
- Johnson, E.J., Goldstein, D. (2003), 'Do defaults save lives?', *Science*, pp. 1338-1339.
- Koszegi, B., Rabin, M. (2007), 'Reference-dependent risk attitudes', *American Economic Review*, vol. 97, pp. 1047-1073.
- Köszegi, B., Rabin, M. (2006), 'A model of reference-dependent preferences', *Quarterly Journal of Economics*, vol. 121, pp. 1133-1166.
- Lien, J.W., Zheng, J. (2015), 'Deciding when to quit: Reference-dependence over slot machine outcomes', *American Economic Review: Papers & Proceedings*, vol. 105, pp. 366-370.
- Loomes, G., Sugden, R. (1986), 'Disappointment and dynamic consistency in choice under uncertainty', *Review of Economic Studies*, vol. 53, pp. 271-282.
- Luce, R.D. (1959), 'On the possible psychophysical laws.', *Psychological Review*, vol. 66, pp. 81.
- Markowitz, H. (1952), 'The Utility of Wealth', *Journal of Political Economy*, vol. 60, pp. 151-158.
- Nilsson, H., Rieskamp, J., Wagenmakers, E. (2011), 'Hierarchical bayesian parameter estimation for cumulative prospect theory', *Journal of Mathematical Psychology*, vol. 55, pp. 84-93.

- Odean, T. (1998), 'Are investors reluctant to realize their losses?', *Journal of Finance*, vol. 53, pp. 1775-1798.
- Pesendorfer, W. (2006), 'Behavioral economics comes of age: A review essay on advances in behavioral economics', *Journal of Economic Literature*, vol. 44, pp. 712-721.
- Pope, D.G., Schweitzer, M.E. (2011), 'Is Tiger Woods loss averse? persistent bias in the face of experience, competition, and high stakes', *American Economic Review*, vol. 101, pp. 129-157.
- Prelec, D. (1998), 'The probability weighting function', *Econometrica*, vol. 66, pp. 497-528.
- Rabin, M. (2000), 'Risk aversion and expected-utility theory: A calibration theorem', *Econometrica*, vol. 68, pp. 1281-1292.
- Rouder, J.N., Lu, J. (2005), 'An introduction to bayesian hierarchical models with an application in the theory of signal detection', *Psychonomic Bulletin & Review*, vol. 12, pp. 573-604.
- Samuelson, W., Zeckhauser, R. (1988), 'Status quo bias in decision making', *Journal of Risk and Uncertainty*, vol. 1, pp. 7-59.
- Schmidt, U., Starmer, C., Sugden, R. (2008), 'Third-generation prospect theory', *Journal of Risk and Uncertainty*, vol. 36, pp. 203-223.
- Stott, H.P. (2006), 'Cumulative prospect theory's functional menagerie', *Journal of Risk and Uncertainty*, vol. 32, pp. 101-130.
- Sugden, R. (2003), 'Reference-dependent subjective expected utility', *Journal of Economic Theory*, vol. 111, pp. 172-191.

- Sydnor, J. (2010), '(Over) insuring modest risks', *American Economic Journal: Applied Economics*, vol. 2, pp. 177-199.
- Thaler, R.H., Benartzi, S. (2004), 'Save more tomorrow™: Using behavioral economics to increase employee saving', *Journal of Political Economy*, vol. 112, pp. S164-S187.
- Toubia, O., Johnson, E., Evgeniou, T., Delquié, P. (2013), 'Dynamic experiments for estimating preferences: An adaptive method of eliciting time and risk parameters', *Management Science*, vol. 59, pp. 613-640.
- Tversky, A., Kahneman, D. (1991), 'Loss aversion in riskless choice: A reference-dependent model', *Quarterly Journal of Economics*, vol. 56, pp. 1039-1061.
- Tversky, A., Kahneman, D. (1992), 'Advances in prospect theory: Cumulative representation of uncertainty', *Journal of Risk and Uncertainty*, vol. 5, pp. 297-323.
- van Osch, S., M.C., van den Hout, W.B., Stiggelbout, A.M. (2006), 'Exploring the reference point in prospect theory: Gambles for length of life', *Medical Decision Making*, vol. 26, pp. 338-346.
- Van Osch, S., Stiggelbout, A.M. (2008), 'The construction of standard gamble utilities', *Health Economics*, vol. 17, pp. 31-40.
- van Osch, S.M.C., Wakker, P.P., van den Hout, W.B., Stiggelbout, A.M. (2004), 'Correcting biases in standard gamble and time tradeoff utilities', *Medical Decision Making*, vol. 24, pp. 511-517.
- Wakker, P.P. (2010), *Prospect Theory: For Risk and Ambiguity*, Cambridge UK: Cambridge University Press.

Wilcox, N. (2012), 'Heterogeneity and Stability of Probability Weights', Plenary lecture at the *FUR conference* in Atlanta, 1st July 2012.

Wu, G., Gonzalez, R. (1996), 'Curvature of the probability weighting function', *Management Science*, vol. 42, pp. 1676-1690.