Formalizing reference dependence and initial wealth in Rabin's calibration theorem

Peter P. Wakker

Econometric Institute, Erasmus University Rotterdam, The Netherlands

October, 2005

Running title: reference dependence

Address for correspondence:
Peter P. Wakker
Econometric Institute
Erasmus University
P.O. Box 1738
3000 DR Rotterdam
The Netherlands
31-(0)10 – 408.12.65 (O)
31-(0)10 – 408.91.62 (F)
e-mail:
Wakker@few.eur.nl
Homepage: http:\www.fee.uva.nl/creed/wakker

Abstract

This paper provides a formalization of reference dependence, initial wealth, and final wealth, concepts that are central in the distinction between classical expected utility and prospect theory. The formalization will clarify some misunderstandings about Rabin’s calibration paradox for expected utility. Cox & Sadiraj (2005) argued that Rabin's paradox can easily be explained in terms of utility of income, which descibes outcomes as changes with respect to a given level and which they consider part of expected utility, and that paradoxes similar to Rabin’s apply to prospect theory and other theories as well. Our formalization shows that utility of income is part of prospect theory and not of expected utility, that utility of income was suggested by Rabin himself as the most plausible explanation of his paradox under the term loss aversion, and that the "similar" paradoxes for prospect theory are, contrary to Rabin's paradox, based on empirically implausible assumptions so that they have no bite.

JEL Classification: D81, C60

Keywords: Rabin’s calibration, reference dependence, utility of income, loss aversion

1. Introduction

This paper formalizes initial wealth, final wealth, and reference dependence in expected utility and prospect theory. The formalization is applied to Rabin’s (2000) calibration paradox for expected utility. Rabin did not formalize the concepts mentioned so as to be maximally accessible to a wide audience. In view of the continued misunderstandings about his paradox, however, a formalization is warranted at this stage. One widespread misunderstanding concerns the utility of income, a term often used for the modeling of outcomes as changes with respect to a given level. Utility of income is often believed to be part of expected utility. That this is not so is demonstrated for instance in Corollary 4.6 below, showing that, for utility of income, risk aversion need not imply concave utility. Utility of income is nothing other than reference dependence of prospect theory, and entails a major breakaway from expected utility.

The formalization of this paper will reveal that the criticism of Rabin's (2000) calibration theorem by Cox & Sadiraj (2005; CS hereafter) is based on misunderstandings as described above. Rabin’s paradox remains a valid descriptive criticism of expected utility. His paradox illustrates particularly clearly that there are many problems with the classical economic modeling of risk attitude through utility curvature.

2. **The innocuous role of initial wealth under expected utility in terms of final wealth**

We first introduce notation. Capital *I*, the *initial wealth*, denotes the value of all assets possessed by an agent prior to a choice now considered. For simplicity, *I* is assumed monetary and deterministic. *Outcomes*, with generic notation x,y, etc. are monetary (real numbers). For each outcome x, the corresponding *final wealth* is *I*+x. Outcomes, thus, designate changes relative to initial wealth. *Prospects* are probability distributions over outcomes. By (p:x, y) we denote the prospect yielding outcome x (final wealth *I*+x) with probability p and outcome y (final wealth *I*+y) with probability 1−p. We equate an outcome x with the prospect of receiving x with certainty. $\succcurlyeq_I$ denotes preferences over prospects given initial wealth *I*, with $\succ_I$ etc. as usual.

Expected utility holds if there exists a utility function U* such that the preference relation $\succcurlyeq_I$ maximizes the expectation of U* over final wealth. For example, the $\succcurlyeq_I$ preference-value of prospect (p:x, y) is pU*(*I*+x) + (1−p)U*(*I*+y). Preferences are not affected if some amount is added to *I* and the same amount is subtracted from all outcomes of all prospects, because such an operation does not affect the final wealth positions involved.

I is considered a characteristic of the agent in the same way as age, gender, etc. are, many of which we do not even know but take as *fixed given the agent*. That is, *I* is usually assumed constant. In this respect *I*, the constant indicating initial wealth, is different than r, the reference point that will be introduced in Section 4. This r will not be constant during the analysis, and preferences will then no longer depend solely on the generated final wealth.

Because it is inconvenient to always denote *I*, and we often do not even know it but only assume that it is constant, *I* is suppressed from the notation in virtually all applications of expected utility.[1] We then write U(x) instead of U*(*I*+x), and $\succcurlyeq$ instead of $\succcurlyeq_I$. Expected utility equals the expectation of U with respect to the prospect when expressed in terms of outcomes. Although outcomes designate a change of wealth and not final wealth, we nevertheless say that *outcomes are in terms of final wealth* in this and the following section, because for every agent every outcome x in our analysis is uniquely related to a final wealth position *I*+x. To emphasize this point, we sometimes say *expected utility in terms of final wealth* instead of the shorter but equivalent expected utility.

Risk aversion indicates preference for the expectation of a prospect over that prospect. The following, trivial, variation of classical results, stated under the usual assumptions such as completeness of preference over the entire domain, is given for the sake of comparison. It illustrates the innocuous role of the constant *I*, and of the substitution of U for U*. Dropping the constant *I*, and replacing U* by U, amounts to nothing more than a convenient rescaling of outcomes.

Observation 2.1. Under expected utility with constant *I*, risk aversion is equivalent to concave U as well as to concave U*. Higher risk premiums correspond to higher values of −U´´/U´ as well as to higher values of −U*´´/U*´. □

---

[1] An exception can be found in parametric fittings of power utility $x^{1-r}/(1-r)$ ("constant relative risk aversion"), where for $r \geq 1$, as commonly found in finance and macroeconomics, utility is not defined at x=0. Then often an extra parameter *I*>0 is introduced, and utility is $(x+I)^{1-r}/(1-r)$. *I* is then often interpreted as initial wealth (Beetsma & Schotman 2001). Other exceptions concern explicit studies of the dependence of risk attitudes on wealth (Guiso & Paiella 2003).

3. **Rabin's paradox for expected utility in terms of final wealth**

The following preference displays (at least) moderate risk aversion in the neighborhood of outcome 0. It is found for most individuals in many samples, at all common levels of initial wealth *I*:

$$0 \succ_I (0.5\!: +11,\ -10)\ . \qquad (3.1)$$

Therefore, the following assumption is plausible.

RABIN'S EMPIRICAL ASSUMPTION. For the common individual, Eq. 3.1 holds not only for their actual level of initial wealth, but it would also hold had the individual been at any other common level of initial wealth. □

For simplicity, we will not formalize what "common levels of initial wealth" are. Any accepted interval of length, say, 5000, of such levels suffices for the following analysis. The first step in establishing Rabin's paradox concerns the derivation of the following preference for "many" nonzero x from his empirical hypothesis:

$$x \succ_I (0.5\!: x+11,\ x-10)\ . \qquad (3.2)$$

Observation 3.1. Under expected utility in terms of final wealth, Rabin's empirical assumption implies that Eq. 3.2 must also be empirically prevailing for many values of x and common levels of initial wealth *I*.

Proof. Consider the choice between x and (0.5:x+11, x−10) as in Eq. 3.2, with nonzero x, for a given agent. If we subtract x from all outcomes of all prospects, and add it to *I*, then preferences should not be affected, leading to the prospects of Eq. 3.1 with, however, initial wealth *I*+x instead of *I*. As long as this level belongs to the common levels, the preference for safety in Eq. 3.1 is prevailing. Hence, so it does in Eq. 3.2. Note how the between-agent assumption of Eq. 3.1 led to Rabin's within-agent empirical assumption, and then to the within-agent assumption in Eq. 3.2. □

If we observe Eq. 3.1 for an individual at some $I \geq 5000$ and add the widely accepted hypothesis of decreasing absolute risk aversion, then Eq. 3.1 holds for this individual at all lower levels of initial wealth, and Eq. 3.2 holds for all x and *I* ranging over [10, 2400] and more. Under expected utility, Eq. 3.2 implies that

$$U(x+11) - U(x) \quad < \quad U(x) - U(x-10). \qquad (3.3)$$

For concave utility it follows that $U'(x+11) \leq \big(U(x+11)-U(x)\big)/11$ (the average marginal utility over the interval [x,x+11]) $\leq$ (because of Eq 3.3) $10/11 \times \big(U(x)-U(x-10)\big)/10$ (the latter fraction is the average marginal utility over the interval [x−10,x]) $\leq 10/11 \times U'(x-10)$. Hence, over the range where Eq. 3.3 holds, U´ drops by a factor of at least 10/11 over each interval of length 21. Over intervals of length 2100, U´ must drop by a factor of at least $(10/11)^{100} = 0.000073$. This geometric decay is too strong, leading to empirically absurd risk aversion. This implication constitutes Rabin’s paradox. At least one of the assumptions that led to it must be empirically invalid. Because Rabin's empirical assumption is empirically convincing, at least one of the assumptions of expected utility in terms of final wealth must be empirically invalid. The next sections consider two such assumptions.

4. **Reference dependence as a fundamental breakaway from the classical paradigm**

Whereas from the perspective of classical expected utility in terms of final wealth there is no real difference between Eq. 3.1 holding for many *I* and Eq. 3.2 holding for many x, and critics of Rabin (2000) often did not distinguish between these claims, the difference will become crucial in this section. Under *reference dependence,* a new parameter r is introduced, the *reference point*, which depends on factors yet to be specified. This new parameter considerably increases the generality of the theory in a manner useful for descriptive purposes. Unfortunately, as a price to pay, it also increases the complexity of the theory. Reference dependence is usually considered not to be normative.

The reference point r is to be distinguished from the initial wealth *I* from the preceding sections. Unlike *I*, r varies within an individual between different choices, and need not be constant during the analysis. The variable r does not serve to capture all assets of the individual as did *I*, but it captures psychological framing heuristics used by the individual. An outcome x now corresponds to final wealth *I*+r+x. The interpretation is that the agent takes r, i.e. final wealth *I*+r, as reference point, and x as change with respect to that reference point. Outcomes here should be distinguished from outcomes in the preceding section. They now designate changes with respect to the reference point, and are no longer uniquely related to final wealth because of the intervening role of the variable r. The following figure illustates the relations between final wealth, r, *I*, and outcomes.

		decomposition of final wealth F	interpretation	evaluation
classical model		F	final wealth	U*(F)
	I constant: innocuous rescaling of outcomes	I + y	initial wealth + outcome	U(y)
reference dependence	r variable: fundamental breakaway from classical model	*I* + r + x	initial wealth + reference point + outcome	U(r,x)

FIGURE. Decompositions of final wealth.

In Section 5, we will consider deviations from expected utility due to probability weighting $p \mapsto w(p)$ for a nonlinear transformation w. To clearly distinguish the analysis of this section from that of Section 5, we display the following assumption, made throughout this section.

Assumption 4.1. The weighting of probabilities is linear as in expected utility. □

We remodel the choice in Eq. 3.2, maintaining the same final wealth, as

$$U(r, x-r) \;>\; 0.5 \times U(r, x+11-r) + 0.5 \times U(r, x-10-r). \qquad (4.1)$$

Now utility U depends on two variables, r and x (besides the constant *I* that is suppressed). Kahneman and Tversky used the symbol v instead of U, and proposed the term value function instead of utility function, to emphasize that their concepts are more general. To stay closer to economic traditions, I will continue to use the term utility function and the symbol U, in the same way as I continue to use the term outcome even though the meaning now is more general. Utility and outcomes of the preceding sections can be considered the special case where r is kept constant (equal to 0), and r is not expressed in notation. Under reference dependence, we do assume that r is the same for all outcomes in one choice situation.

Loss aversion entails, loosely speaking, that $U(r,x)$ as a function of its second argument is steeper for negative arguments x (*losses*) than for positive ones (*gains*). It can be interpreted as extreme concavity of utility at $x = 0$, with a nondifferentiable kink there. A moderate degree of loss aversion suffices to explain Rabin's empirical assumption under the plausible assumption of $r = 0$ there, so that the reference point corresponds to initial wealth in that choice situation.

Observation 4.2. The derivation in Observation 3.1 of Eq. 3.2 from Rabin's empirical assumption fails under reference dependence and loss aversion.

Explanation. In the proof of Observation 3.1, it is plausible that the agent at initial wealth I+x instead of I, when faced with the two options in Eq. 3.1, continues to have $r = 0$, so that the reference point then corresponds to final wealth I+x and not to final wealth I as in Eq. 3.2. The outcomes, i.e. changes with respect to the reference point, will then be 0 for the safe option, and 11 and –10 for the risky option, again different from the changes x, x+11, and x–10 in Eq. 3.2. The choices in the two situations concern different reference points and different outcomes and, obviously, need not agree. The safe choice in Eq. 3.1 no longer implies the safe choice in Eq. 3.2. □

The following example, similar to Sections 2.2 and 3.1 of CS, illustrates the above observation.

Example 4.3. Take utility U(r,x) = r+x for $x \geq 0$ and U(r,x) = r + λx for $x \leq 0$ with $\lambda > 11/10$, the latter consistent with the common empirical findings of $\lambda \approx 2$. Then Eq. 3.1 in Rabin's empirical assumption is satisfied but Eq. 4.1 and, hence, Eq. 3.2 are violated for all $x \geq 10$ (avoiding losses) when always r=0 is taken in these choices. Rabin's empirical assumption Eq. 3.1 does not imply Eqs. 3.2 and 4.1. □

We conclude that Rabin's paradox can be accommodated by loss aversion, without implying extreme behavior of U. This reasoning formalizes the argument in the last paragraph of Rabin (2000). It accommodates Eq. 3.1, but not Eq. 3.2.[2] Hence, the explanation of Rabin's paradox solely in terms of loss aversion needs the reasoning of Observation 3.1 to generate the expected-utility paradox. The following observation is trivial, but is presented for the sake of comparison.

Observation 4.4. Let Assumption 4.1 hold. Then for choices restricted to any fixed reference point r, with further the usual assumptions such as completeness of preference over the entire domain, the classical results of expected utility remain true, with risk aversion equivalent to a concave utility function U(r,·), with higher risk premia corresponding to higher values of $-U(r,\cdot)''/U(r,\cdot)'$, etc. □

---

[2] If we assume that $U^*(I+x) = U(x) = x^{0.88}$ and r = 0, then Eq. 3.2 is satisfied for $x \leq 14$ but violated for $x \geq 15$.

The increased generality of r, obviously, shows up only if we consider variations of r. The following theorem illustrates this increased generality. The variations in r then reveal that the usual assumption of completeness of preference over the entire domain of prospects can easily be violated. Consequently, risk aversion need no longer imply concave utility. This finding illustrates once more that reference dependence with a variable parameter r entails a fundamental breakaway from expected utility. It also illustrates the different role for r than for the constant parameter *I* (compare Observation 2.1).

Theorem 4.5 [Risk Aversion Explained by Loss Aversion with possibly nonconcave utility]. Let Assumption 4.1 hold. Assume that an agent, whenever choosing between a risky prospect and a sure amount corresponding to final wealth *I*+y, takes reference point r = y. Assume also that utility is always steeper for losses than for gains, i.e. $U'(r,\ell) \geq U'(r,g)$ for all r and all $\ell < 0 < g$ (derivative with respect to the second argument). Then risk aversion holds, i.e. each prospect is less preferred than its expectation.

Proof. For the ordering between a sure outcome and the expectation of a prospect and, hence, for the definition of risk aversion, it does not matter whether we describe outcomes in terms of final wealth, final wealth minus initial wealth *I*, or final wealth minus *I*+r for a reference point r. Consider a probability distribution P* over final wealth, and assume that $I+y \geq \int_{\mathbb{R}} x dP^*$ for some outcome y. That is, *I*+y exceeds the expectation of P*. To demonstrate that risk aversion holds, we have to show that the sure outcome corresponding to final wealth *I*+y is preferred to the prospect corresponding to P*.

To apply reference dependent theories, we first have to specify the reference point r. For each r, define μ_r such that $U'(r,\ell) \geq \mu_r \geq U'(r,g) \geq 0$ for all $\ell < 0 < g$, and set $U(r,0) = 0$. In this theorem, it is assumed that the reference point r is y (final wealth I+r = I+y). We write r instead of y henceforth. Let P be the probability distribution resulting from P* by subtracting I+r from all final wealth levels, i.e., it is the probability distribution over outcomes (changes with respect to the reference point I+r) corresponding to P*. Because $I + y \geq \int_{\mathbb{R}} x dP^*$, we have $0 \geq \int_{\mathbb{R}} x dP$.

The reference-dependent evaluation of P is $\int_{\mathbb{R}} U(r,x)dP = \int_{\mathbb{R}^+} U(r,x)dP + \int_{\mathbb{R}^-} U(r,x)dP \leq \int_{\mathbb{R}^+} \mu_r x dP + \int_{\mathbb{R}^-} \mu_r x dP = \mu_r \int_{\mathbb{R}} x dP \leq \mu_r 0 \leq 0 = U(r,0)$. The reference point is preferred to the prospect, and risk aversion follows. □

The choice of reference point in the above theorem is psychologically plausible (Herschey & Schoemaker 1985; Johnson & Schkade 1989; Robinson, Loomes, & Jones-Lee 2001; van Osch & Stiggelbout 2005). The assumption of steeper utility for losses than for gains is, obviously, less restrictive than concavity of utility. It was proposed as a formal definition of loss aversion by Bowman, Minehart, & Rabin (1999). The theorem sheds new light on the richness assumption in Observation 4.4, of preferences between all prospect pairs being observable from all reference points r. Even though this assumption is commonly made in theoretical papers on prospect theory, it is less convincing than richness assumptions usually are. This point was raised by Bleichrodt (2005), who provided the first theoretical analysis to relax this assumption, involving advanced mathematical derivations.

Theorem 4.5 suggests another empirical point. Probably most of the risk aversion empirically observed is generated by loss aversion. Because expected utility ignored this aspect, it had to use utility to model risk aversion. The sharp kink and concavity at a point when being a reference point as in Theorem 4.5, was lumped together with what in fact is close to linearity when that point is not a reference point, leading to overly concave utility functions. These overly concave utility functions were "misused" to accommodate the risk aversion that is in fact generated by loss aversion. Empirical evidence supporting this point can be found in Abdellaoui, Bleichrodt, & Paraschiv (2004). The following corollary of Theorem 4.5 illustrates this point for expected utility.

Corollary 4.6. Let Assumption 4.1 hold. Under reference dependence, risk aversion can hold with utility U(r,x) strictly convex both for gains $x > 0$ and for losses $x < 0$. □

Given the generality of reference dependence, with an extra parameter r, the new phenomena of Theorem 4.5 and Corollary 4.6 should come as no surprise. We discuss this point further in Section 7.

5. **Probability weighting as alternative explanation of Eqs. 3.1 and 3.2**

As indicated by CS (Section 4.1) and Rabin (2000, penultimate paragraph of main text), Eq. 3.1 and, in fact, all of Eq. 3.2, can be explained by probability weighting without reference dependence. To illustrate this point we assume, relative to Section 4, that r is 0 throughout this section, and do not denote it in the argument of U. Eq. 3.2 then amounts to

$$U(x) > w(0.5)U(x+11) + (1-w(0.5))U(x-10), \tag{5.1}$$

where w is the probability weighting function.[3] Under the common parameters found for prospect theory (Abdellaoui 2000; Bleichrodt & Pinto 2000; Gonzalez & Wu 1999; Tversky & Kahneman 1992), w(0.5) = 0.42 on average. This estimate, together with any linear or concave utility function, accommodates Eq. 5.1 and, thus, Eqs. 3.2 and 3.1. Again, no extreme behavior of U and no paradoxes are implied.

Because Eq. 3.2 is empirically plausible in its own right, at least for gains, a variation of Rabin's paradox could be devised where not Eq. 3.1, but all of Eq. 3.2 is taken as the empirical assumption, and then the absurd behavior of utility is derived as before. This variation of the paradox does not need the reasoning of Observation 3.1.

The first part of this section has demonstrated, in agreement with CS, Section 4.1 and with Rabin (2000, penultimate paragraph of main text), that loss aversion is not the only factor that can explain Eq. 3.1 and Rabin's paradox, and that expected utility has more problems. I nevertheless agree with Rabin & Thaler (2001) that loss aversion, while not the only, is still the main factor underlying Rabin's paradox. Accordingly, I prefer the version of the paradox presented in Section 3, with only Eq. 3.1 as empirical assumption, to the variation described above. Whereas Eq. 3.2 is already plausible, Eq. 3.1 is considerably more plausible. Loss aversion generates considerably more of the risk aversion in Eq. 3.1 than probability weighting does. Loss aversion is one of the strongest phenomena in the field of risky choice.

Other criticisms of Rabin (2000), by Watt (2002) and Palacios-Huerta & Serrano (2005), considered the above variation of Rabin's paradox, with Eq. 3.2 instead of Eq. 3.1 as empirical assumption. They criticized the empirical plausibility of Eq. 3.2, not

[3] CS denote w as h and apply it to the worst outcome of the prospect, as was common in the first papers on rank dependence. We use the dual notation where w is applied to the best outcome, as is more popular today.

on the basis of these choices, but on predictions regarding these choices derived from other data in the literature *while assuming expected utility*. They, thus, did not consider Eq. 3.1 and Observation 3.1, nor the empirical plausibility of Eq. 3.2 for many x on the basis of deviations from expected utility such as probability weighting.

CS (Section 4.2) pointed out that the extreme behavior of the utility function is also implied under probability weighting if Eq. 3.2 can be replaced by

$$x \succ (p{:}x+11, x-10) \qquad (5.2)$$

with p such that w(p) = 0.5 (which happens on average for p approximately 0.64), and if it can be assumed that this equation holds for many x. The same argument was advanced before by Neilson (2001). The algebraic derivation is identical to that of Rabin for expected utility and is, obviously, theoretically correct.

Eq. 5.2 does not entail a paradox though, contrary to Rabin's finding, because the premise of Eq. 5.2 holding for many x is not empirically plausible, contrary to Rabin's empirical assumption of Eq. 3.1. Eq. 5.2 requires considerably more risk aversion. Under the parametric estimations of Tversky & Kahneman (1992), with $U(x) = U^*(I+x) = x^{0.88}$, Eq. 5.2 is violated for all $x \geq 15$, and common descriptive theories do not predict it, contrary to Rabin's Eq. 3.1. It is not informative to derive implausible implications for utility from implausible empirical assumptions. CS similarly analyze preferences $x \succcurlyeq (0.5{:}x-75, x+110)$ for all $x \geq 75$, but such preferences are not plausible for large x either and, for instance, are not predicted by Tversky & Kahneman (1992) for $x \geq 175$.

6. **Utility of income**

Reference dependence has often been used in experimental economics, and right so given its descriptive realism. It is then usually referred to as utility of income, for instance in auction experiments. Unfortunately, it has sometimes been suggested that utility of income is a minor variation of expected utility, and is best headed under the expected utility models. One argument advanced is that expected utility when taken as an abstract mathematical theory does not speak to the nature of outcomes. It then does not specify whether outcomes are final wealth, are uniquely related to final wealth, are different for odd minutes on a day than for even ones, are different when in the left hand than when in the right hand, are different when in your pocket than when in your hand, etc.

The different ways of modeling outcomes just described may all be equally interesting to mathematicians, but they are not to economists. Economics is not an abstract mathematical theory, but is about human beings and money, and there are agreed-upon conventions of modeling. Economists are not interested in a currency that changes from odd to even minutes on a day in the same way as ornithologists are not interested in so-called blite ravens. Blite ravens have been black up to five minutes ago and are white thereafter, and were discussed in studies of inductive reasoning (Goodman 1965). I, and many economists alike, consider expected utility to be rational if outcomes are monetary in terms of final wealth, but not if outcomes are twice the dollars in your right hand plus once your dollars outside your right hand, and neither if outcomes are changes with respect to a (nonfixed) reference point.

In particular descriptive circumstances, r=$5 in the pocket and x=$15 in the right hand can be perceived differently than r=$10 in the pocket and x=$10 in the right hand, and then special descriptive theories to distinguish can be useful. Thaler (1996) considered particular kinds of such theories, and Luce (2000, Chapters 4−7) developed advanced mathematics for this phenomenon. Reference dependence and utility of income belong to such descriptive theories. They imply descriptive phenomena markedly different than those predicted by expected utility (Corollary 4.6).

The normative status of the descriptive theories just mentioned is very different from that of expected utility in terms of final wealth. I hope and pray that the dear term expected utility, commonly used to designate the Bayesian hallmark of rationality, will not be confused with something as irrational as reference-dependence/utility-of-income.[4] I conjecture that several misunderstandings in the literature about this difference result from confusions between the innocuous role of the constant initial wealth *I* versus the crucial role of the variable r, and confusions between the classical Observations 2.1 and 4.4 versus the new phenomenon in Corollary 4.6.

[4] There are good reasons to believe that rational behavior should be close to risk neutrality for moderate stakes. Then reference dependence does not affect behavior and is equivalent to decisions in terms of final wealth, so that it does not entail irrationalities. It can then facilitate calculations (bounded rationality). The main text only refers to cases where reference dependence essentially affects behavior and essentially deviates from decisions in terms of final wealth, and where bounded rationality plays no role.

7. **History and Applications of Reference Dependence**

Because there have been misunderstandings about the novelty of utility of income and its relation to reference dependence, we discuss the history and current status of these concepts. Reference dependence has been around for a long time. Usually, Markowitz (1952) is credited as the first to have proposed this phenomenon clearly. On p. 157, he immediately pointed out that the absence of a theory about the location of r is problematic. Edwards (1954) discussed the phenomenon extensively (p. 395, 400). The earliest statement of loss aversion that I am aware of is in Robertson (1915, p. 135). Pfanzagl (1959, p. 290) used the expression “the amount of money in front of the subject” to designate x and “the amount in his pocket” to designate the reference point r. Arrow (1951, p. 432) discussed early proposals and criticized them for the absence of a theoretical basis for the choice of a reference point (“zero point”). Reference dependence is half of the way in which prospect theory breaks away from expected utility, with nonlinear probability weighting the other half (Kahneman & Tversky 1979, Tversky & Kahneman 1992).

Many empirical studies have suggested that loss aversion is one of the most prominent empirical phenomena in decision theory. Hence, there is much interest in reference dependence and loss aversion, in spite of several theoretical difficulties. Reference dependence and loss aversion depend much on details of framing. There is common agreement that they can be reduced and even, hopefully (given their irrational nature), can be removed under proper explanations, learning, and motivation (Bleichrodt, Pinto, & Wakker 2001; Payne, Bettman, & Schkade 1999; Plott & Zeiler 2005; Tversky & Kahneman 1986 p. S273).

Without further specification of the location of r, of its dependence on decision contexts, and of the dependence of U on r, reference dependence is too general to yield predictions. In most mathematical and axiomatic studies it is assumed that the location of r has already been determined, and r is taken fixed (Tversky & Kahneman 1992). Then the choice of r is part of the modeling stage that precedes the formal analysis, based on heuristics and what is known today.

There have been several psychological investigations into reference dependence, with examinations of eye movements (Johnson & Schkade 1989) and of speak-aloud protocols (Lopes 1987; Robinson, Loomes, & Jones-Lee 2001; van Osch & Stiggelbout 2005). Only recently, theoretical studies have begun to consider variations of r (Bleichrodt 2005; Köszegi & Rabin 2005; Schmidt 2003; Schmidt, Starmer, & Sugden 2005).

Prior to any application of reference dependence, plausible assumptions have to be made about the reference point. For example, in auction theory it is usually assumed implicitly that the reference point corresponds to the situation of the subjects immediately prior to the auctions, and that $U(r,\cdot)$ is independent of r. This is a good example of a clear and plausible assumption that is specific enough to make the theory operational and tractable.

8. **Cox & Sadiraj (2005)**

CS (and, similarly, Rubinstein 2002) plead for utility of income as the most plausible explanation of Rabin's paradox. As far as I can see, their explanation is the same as Rabin's explanation of reference dependence (plus loss aversion). Yet, CS seem to distinguish between these explanations. In many places they (and, similarly, Rubinstein 2002) suggest that the utility-of-income model belongs to expected utility (e.g., end of 1st para of introduction), and that it involves no more than a re-interpretation of outcomes. I argued differently above. CS further give some results similar to our Observation 4.4 in their Section 3.2.

CS also seem to assume that loss aversion does not speak to the problems discussed above because they claim that all outcomes involved can be restricted to gains (end of Section 1; p. 19, Concluding Remarks). Apparently, for final wealth I + r + x, they do not let the sign and gain-or-loss status be determined by x (which is a negative $-\ell$ in their Eq. 4, called “loss amount” there), but by r+x or, possibly, I+r+x. That is, they misunderstood the terminology of prospect theory.

Although CS describe outcomes as “income” or “changes in wealth” (their Section 2.2), they never specify what the reference point is relative to which these changes are to be taken. In particular, it cannot be inferred from their paper how this reference point might differ from the reference point of prospect theory. Any such difference would be highy implausible.

When claiming that Eq. 5.2 is equally problematic (or nonproblematic) for prospect theory as Eq. 3.2 is for expected utility, CS, strangely enough, write several times (p. 5 l. 6, end of section 1; p. 14, beginning of Section 4; Section 5, Concluding Comments) that their paper will not discuss the empirical plausibility of these equations. Neilson (2001) was, similarly, silent on the empirical plausibility of these equations. As argued above, this plausibility is crucial and cannot be ignored. Deriving implausible utilities from implausible assumptions is not informative.

9. **Conclusion**

Rabin's empirical assumption, which leads to a paradox for expected utility, can be explained by the parameter estimations of prospect theory (Tversky & Kahneman 1992) and, even stronger, can be explained by loss aversion alone (Section 4) and also by probability weighting alone (Section 5). No extreme risk aversion for large stakes, and no paradoxes, follow under these explanations.

The crucial novelty of reference dependence is not that outcomes are perceived as changes rather than absolute amounts, but rather that the comparison-level for the changes is not constant (initial wealth) but is variable (initial wealth plus r) during the analysis.

In the end it is not important who was first on what, and how ideas are called. Important is that the right ideas survive. The reference-dependent/income-utility way of modeling outcomes is descriptively realistic, and different fields from different perspectives, including prospect theory and auction theory, are converging to it.

Acknowledgment.

Han Bleichrodt made helpful comments.

References

Abdellaoui, Mohammed, Han Bleichrodt, & Corina Paraschiv (2004), “Measuring Loss Aversion under Prospect Theory: A Parameter-Free Approach,” iMTA/iBMG, Erasmus University, Rotterdam, The Netherlands.

Arrow, Kenneth J. (1951), “Alternative Approaches to the Theory of Choice in Risk-Taking Situations,” *Econometrica* 19, 404–437.

Beetsma, Roel M.W.J. & Peter C. Schotman (2001), “Measuring Risk Attitudes in a Natural Experiment: An Empirical Analysis of the Television Game Show LINGO,” *Economic Journal* 111, 821–848.

Bleichrodt, Han (2005), “Reference-Dependent Utility without Completeness,” iMTA, Erasmus University, Rotterdam, The Netherlands.

Bleichrodt, Han, José Luis Pinto, & Peter P. Wakker (2001), “Making Descriptive Use of Prospect Theory to Improve the Prescriptive Use of Expected Utility,” *Management Science* 47, 1498–1514.

Bowman, David, Deborah Minehart, & Matthew Rabin (1999), “Loss Aversion in a Consumption-Savings Model,” *Journal of Economic Behavior and Organization* 38, 155–178.

Cox, James C. & Vjollca Sadiraj (2005), “Small- and Large-Stakes Risk Aversion: Implications of Concavity Calibration for Decision Theory,” *Games and Economic Behavior*, forthcoming; undated version downloaded from http://www.econlab.arizona.edu/~jcox/working_papers/ConcavityCalibration.pdf on July 2, 2005.

Edwards, Ward (1954), “The Theory of Decision Making,” *Psychological Bulletin* 51, 380–417.

Gonzalez, Richard & George Wu (1999), “On the Shape of the Probability Weighting Function,” *Cognitive Psychology* 38, 129–166.

Goodman, Nelson (1965), “Fact, Fiction and Forecast.” Bobbs-Merrill, Indianapolis.

Guiso, Luigi & Monica Paiella (2003), “Risk Aversion, Wealth and Background Risk,” Bank of Italy Economic Working Paper No. 483.

Hershey, John C. & Paul J.H. Schoemaker (1985), “Probability versus Certainty Equivalence Methods in Utility Measurement: Are They Equivalent?,” *Management Science* 31, 1213–1231.

Johnson, Eric J. & David A. Schkade (1989), “Bias in Utility Assessments: Further Evidence and Explanations,” *Management Science* 35, 406–424.

Kahneman, Daniel & Amos Tversky (1979), “Prospect Theory: An Analysis of Decision under Risk,” *Econometrica* 47, 263–291.

Köszegi, Botond & Matthew Rabin (2005), “A Model of Reference-Dependent Preferences.”

Lopes, Lola L. (1987), “Between Hope and Fear: The Psychology of Risk,” *Advances in Experimental Psychology* 20, 255–295.

Luce, R. Duncan (2000), “*Utility of Gains and Losses: Measurement-Theoretical and Experimental Approaches.”* Lawrence Erlbaum Publishers, London.

Markowitz, Harry M. (1952), “The Utility of Wealth,” *Journal of Political Economy* 60, 151–158.

Neilson, Wiliam S. (2001), “Calibration Results for Rank-Dependent Expected Utility,” *Economics Bulletin* 4, 1–5.

Palacios-Huerta, Ignacio & Roberto Serrano (2005), “Rejecting Small Gambles under Expected Utility,” Dept. of Economics, Brown University, Providence.

Payne, John W., James R. Bettman, & David A. Schkade (1999), “Measuring Constructed Preferences: Towards a Building Code,” *Journal of Risk and Uncertainty* 19, 243–270.

Pfanzagl, Johann (1959), “A General Theory of Measurement —Applications to Utility,” *Naval Research Logistics Quarterly* 6, 283–294.

Plott, Charles R. & Kathryn Zeiler (2005), “The Willingness to Pay–Willingness to Accept Gap, the “Endowment Effect,” Subject Misconceptions, and Experimental Procedures for Eliciting Valuations,” *American Economic Review* 95, 530–545.

Rabin, Matthew (2000), “Risk Aversion and Expected-utility Theory: A Calibration Theorem,” *Econometrica* 68, 1281–1292.

Rabin, Matthew & Richard H. Thaler (2001), “Anomalies: Risk Aversion,” *Journal of Economic Perspectives* 15, 219–232.

Robertson, Dennis H. (1954), “Utility and All What?,” *Economic Journal* 64, 665–678.

Robertson, Dennis H. (1915), “*A Study of Industrial Fluctuation; An Enquiry into the Character and Causes of the So-Called Cyclical Movement of Trade*.” P.S. King & Son ltd., London.

Robinson, Angela, Graham Loomes, & Michael Jones-Lee (2001), “Visual Analog Scales, Standard Gambles, and Relative Risk Aversion,” *Medical Decision Making* 21, 17–21.

Rubinstein, Ariel (2002), “Comments on the Risk and Time Preference in Economics,” text of lecture on Dec. 5th, 2001.

Schmidt, Ulrich (2003), “Reference Dependence in Cumulative Prospect Theory,” *Journal of Mathematical Psychology* 47, 122–131.

Schmidt, Ulrich, Chris Starmer, & Robert F. Sugden (2005), “Explaining Preference Reversal with Third-Generation Prospect Theory.”

Thaler, Richard H. (1996), “*Quasi Rational Economics.*” Russell Sage Foundation, New York.

Tversky, Amos & Daniel Kahneman (1992), “Advances in Prospect Theory: Cumulative Representation of Uncertainty,” *Journal of Risk and Uncertainty* 5, 297–323.

Tversky, Amos & Daniel Kahneman (1986), “Rational Choice and the Framing of Decisions,” *Journal of Business* 59, S251–S278.

van Osch, Sylvie M.C., & Anne M. Stiggelbout (2005), “The Construction of Standard Gamble Utilities,” Leiden University Medical Center, Leiden, The Netherlands.

Watt, Richard (2002), “Defending Expected Utility Theory,” *Journal of Economic Perspectives* 16, 227–228.