



Probability weighting for losses and for gains among smallholder farmers in Uganda

Arjan Verschoor¹  · Ben D'Exelle¹

Accepted: 28 November 2020
© The Author(s) 2020

Abstract

Probability weighting is a marked feature of decision-making under risk. For poor people in rural areas of developing countries, how probabilities are evaluated matters for livelihoods decisions, especially the probabilities associated with losses. Previous studies of risky choice among poor people in developing countries seldom consider losses and do not offer a refined tracking of the probability-weighting function (PWF). We investigate probability weighting among smallholder farmers in Uganda, separately for losses and for gains, using a method (common consequence ladders) that allows refined tracking of the PWF for a population with low levels of literacy. For losses, we find marked probability weighting near zero, which is in line with evidence found in Western labs. For gains, the absence of probability weighting is remarkable, particularly its absence near 100%. We also find marked differences in probability weighting for traditional farmers which are in line with the observed livelihoods strategies in the study area.

Keywords Cumulative prospect theory · Probability weighting · Losses · Gains · Uganda

1 Introduction

How do poor people in the rural areas of developing countries, whose livelihoods are subject to the vagaries of nature, thin or absent markets, dysfunctional institutions and erratic governments, evaluate risky prospects? Contrary to a stereotype of what keeps them poor, smallholder farmers in developing countries may exhibit remarkable tolerance of risk, and are considerably less risk averse than

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11238-020-09796-8>) contains supplementary material, which is available to authorized users.

✉ Arjan Verschoor
a.verschoor@uea.ac.uk

¹ School of International Development, University of East Anglia, East Anglia, UK

typical Western populations (Vieider et al. 2019). In this study, we investigate, for a sample of smallholder farmers from a developing country, not risk aversion as such but a different aspect of decision-making under risk with important implications for livelihoods decisions: probability weighting.

In risky choice decisions, how probabilities are evaluated is crucial. Economists typically think of this as probabilities of outcomes being transformed into decision weights on those outcomes, through some psychologically plausible nonlinear probability-weighting process. Following Allais (1953), violations of the independence axiom of expected utility theory have been observed in a large number of experimental studies, which points to the prevalence of nonlinear probability weighting.¹ The rationale for probability weighting that is currently most commonly invoked is due to prospect theory (Kahneman and Tversky 1979; Tversky and Kahneman 1992). In prospect theory, the so-called “diminishing sensitivity” relative to a reference point is the central organising principle, which governs both the distinct shape of the value function and the shape of the probability weighting function (PWF). In the context of probability weighting, this means that people are less sensitive to changes in probabilities that take place further from a reference probability. For instance, if the reference probability is 0, then weighted probabilities will increase by more when a probability of an outcome changes from 5 to 6% than when it changes from 10 to 11%.

Typically, the reference probabilities 0 and 1 are jointly postulated: the certainty that an outcome will not occur and the certainty that it will occur. Diminishing sensitivity relative to these two reference probabilities produces what Fehr-Duda and Epper (2012, p. 569) call the “famous inverse S”: a PWF that is steep near the reference probabilities of 0 and 100%, and relatively flat in the middle, giving rise to the overweighting of small probabilities and the underweighting of large probabilities (Kahneman and Tversky 1979; Tversky and Kahneman 1992; Wu and Gonzalez 1996; Gonzalez and Wu 1999; Starmer 2000; Takahashi 2011). The inverse S-shaped PWF has been confirmed in the bulk of the large number of experimental studies of probability weighting.² The robustness of this finding suggests that diminishing sensitivity relative to the reference probabilities of 0 and 1 provides a good account of how humans transform probabilities into decision weights.

In this study, we investigated as thoroughly as we deemed feasible the probability-weighting habits of small-scale farmers in a poor country. We were motivated by three considerations. First, these farmers are “experts” on decision-making under risk. Frequent hazards such as droughts, floods, pests, and diseases affect their investments, and getting investment decisions right is a matter of life and death (Fafchamps 2003). To the extent that probability weighting is a bias that

¹ Examples include the common consequence effect tests cited in Wu and Gonzalez (1998, pp. 131–2) and the common consequence and common ratio effect tests reviewed in Starmer (2000) and in Fehr-Duda and Epper (2012).

² See the reviews of these studies in Wu and Gonzalez (1996, 1998), Prelec (1998), Gonzalez and Wu (1999), Starmer (2000), Sugden (2004), Stott (2006), Van de Kuilen and Wakker (2011), and Fehr-Duda and Epper (2012). Van de Kuilen and Wakker (2011)’s Online Appendix G lists 50 references to studies that find weighting functions that exhibit the inverse-S shape. Studies that report exceptions to the common finding of an inverse S-shaped PWF are listed in Blavatskyy (2006).

experienced decision-makers should be less affected by, nonlinear probability weighting should be less pronounced for such farmers than for typical experimental subjects. Second, most of the experimental literature on nonlinear probability weighting does not obtain a separate PWF for gains and losses; to the best of our knowledge, Vieider et al. (2019) is the only previous experimental study of risky choice in a developing country that obtains a *separate* PWF for losses. Since hazards are so common in rural areas of developing countries, losses are a frequent occurrence. Knowing the PWF for losses is therefore important, and we investigate it in this study as well as for gains. Third, we were motivated by the intriguing possibility that, in nonstandard subject pools, probability weighting could be very different from that typically found in Western labs. In particular, we wondered whether the all-pervasive nature of risk in the rural areas of developing countries would lead to the emergence of reference probabilities not equal to but somewhere between 0 and 1. If certainty is at best an abstract concept for a population, then perhaps more realistic (historically informed) probabilities would act as referents. If so, then the PWF may no longer resemble an inverse S.

For investigating probability weighting, we designed and implemented common consequence effect tests (Wu and Gonzalez 1998). There are two reasons that we decided to choose this method. The first is that we needed a simple method for a subject pool with low levels of literacy and numeracy (Dave et al. 2010). Humphrey and Verschoor (2004a, 2b) had previously successfully implemented common consequence effect tests in similar subject pools to ours. Based on extensive piloting, we found that we could considerably increase the refinement of these tests while maintaining excellent subject comprehension. Conceptually attractive alternatives such as Van de Kuilen and Wakker (2011)'s midweight method were deemed to be cognitively too demanding for our subject pool. The second is that we needed a method with sufficient tracking ability of the PWF to investigate whether more than two reference probabilities exist (e.g., 0, 1, and one or more others). For that reason, we decided not to make use of functional forms of the PWF, since these rule out *a priori* plausible shapes influenced by reference probabilities at the extremes as well as in between those extremes of the domain of the PWF (as explained below).

Each subject faced ten choices between two three-outcome lotteries. A pair of lotteries may be thought of as representing a “rung” on a “common consequence ladder”. The rungs are related to each other through an identical manipulation of both lotteries, a so-called common consequence shift. In our case, the manipulation consists in shifting identical probability mass from the worst outcome to the intermediate outcome of both lotteries on one rung, which yields the lotteries on another rung. A preference reversal violates the independence axiom of expected utility theory. Assuming cumulative prospect theory (Tversky and Kahneman 1992),

preference reversals permit pronouncements on the relative steepness of the PWF in precisely defined probability intervals.³

Importantly, we implemented two game conditions, in a between-subject design. One condition is in the domain of losses, the other in the domain of gains. In the losses version of the experiment, the best outcome of the three-outcome lotteries is equal to the neutral outcome, so that the intermediate outcome and worst outcome represent losses. In the gains version, the worst outcome is equal to the neutral outcome, so that the intermediate outcome and best outcome represent gains. We established the neutral outcome by giving each subject a voucher 3 weeks before the day of the experimental session in which that subject would participate. The voucher showed the name, address, and portrait photo of the face of the subject, as well as the figure of 8000 shillings prominently displayed. 8000 Ugandan shillings are about twice the median daily wage in the study area, in which waged labour is moreover hard to come by. In a scripted, orally delivered message when the vouchers were handed over, subjects were informed that, depending on the decisions they would be asked to take in three weeks' time, their final earnings could be higher or lower than 8000 shillings.

By making the figure of 8000 shillings salient in this way, we aimed to establish this amount as a reference point, so that lower amounts would be thought of as losses, and higher amounts as gains. The reason we handed over the vouchers 3 weeks before experimental days was to alleviate concerns about a house-money effect (Thaler and Johnson 1990) through inducing a sense of entitlement to the 8,000 shillings through the passage of time.

In the domain of losses, our investigations focused on probabilities between 0 and 0.8. This is mainly for the sake of realism: investment prospects in which losses are more than 80% likely would not be considered by small-scale farmers in a poor country. We find that probability weighting for losses is pronounced and consistent with diminishing sensitivity relative to the reference probability of 0. Weak evidence is obtained for a second reference probability of 0.5, but we deem it to be insufficient for basing a firm conclusion on.

In the domain of gains, we investigated the curvature of the PWF for probabilities between 0.35 and 1. This is again for realism in that investments that have a success probability lower than 35% would not be contemplated in real life. We find no evidence for probability weighting near the probability of 1, so no evidence that 100% acts as a reference probability relative to which diminishing sensitivity is at work. We do find evidence for a PWF that is relatively flat in the middle, which is consistent with an inverse S shape. However, more striking than that is the overwhelming support for expected utility theory (EUT): 42 out of a total of 45

³ The vast majority of experimental studies of probability weighting assume either cumulative prospect theory (Tversky and Kahneman 1992) or rank-dependent utility theory (Quiggin 1982). If all outcomes are neutral or gains, then the two theories are indistinguishable. However, rank-dependent utility theory has no place for losses relative to a neutral outcome, which is why we assume cumulative prospect theory. Compared to Kahneman and Tversky (1979)'s original prospect theory, it is an improvement, because it corrects the original theory's problem that stochastically dominated lotteries may be sometimes preferred. See Wakker (2010) for more elaborate discussion of these theories. The more recent influential theoretical development of reference-dependent preferences of Koszegi and Rabin (2006, 2007) abstracts from probability weighting, which makes it unsuitable for our purposes.

common consequence steps (comparisons between rungs) are consistent with EUT and only 3 point to nonlinear probability weighting.

When investigating correlates of probability weighting, we find that traditional farmers, those who farm mainly for subsistence and with a minimal reliance on purchased inputs, evaluate probabilities differently from the rest of the sample. Whereas concavity near zero characterises the PWF for the sample as a whole in the domain of losses, convexity does so for traditional farmers. In the domain of gains, whereas linearity of the PWF when approaching 100% characterises the aggregate sample, convexity does so for traditional farmers. We argue in the paper that the contrast between how traditional and nontraditional farmers evaluate probabilities makes sense of the observed difference between them in livelihoods strategies.

The main lesson from our study combines our findings from the domains of losses and gains. In a sample of experienced decision-makers under risk, risky choice in the domain of losses points to pronounced probability weighting that is largely consistent with experimental evidence from Western labs, whereas risky choice in the domain of gains is largely consistent with EUT. Of particular relevance to this special issue, we argue that our results suggest that poor farmers in developing countries should not be seen as shying away from risk but as particularly adept decision-makers under risk (cf. Vieider et al. 2019).

We see the contribution of our study as follows. To begin with, we contribute to the literature on individual risky choice in developing countries. Previous studies have investigated probability weighting (Humphrey and Verschoor 2004a, b; Harrison et al. 2010; Tanaka et al. 2010; Liu 2013; l'Haridon and Vieider 2018; Vieider et al. 2018), but none of these have tracked the PWF with the refinement offered here. Humphrey and Verschoor (2004a, 2004b) use only the probabilities 0, 1/4, 1/2, 3/4, and 1. In other studies of probability weighting in nonstudent samples in developing countries, the parameters of a functional form are estimated as those most likely to generate the experimental data given some assumptions on errors; for instance, Prelec (1998)'s two-parameter functional form of the PWF is estimated both in Harrison et al. (2010) and in Vieider et al. (2018). In the design of Tanaka et al. (2010), which they used in rural Vietnam and Liu (2013) used in rural China, the parameter of Prelec (1998)'s one-parameter functional form of the PWF is instead obtained directly based on a series of paired lottery choices. Shapes with more than one convex or more than one concave area cannot be captured by these functional forms. For that reason, they cannot capture shapes influenced by reference probabilities at the extremes as well as in between those extremes of the domain of the PWF. We deem such shapes plausible in our subject pool; see Sect. 2.2 for further discussion.⁴ Moreover, imposing a functional form may hide linearity of the PWF for ranges of its domain for which decision-makers do not transform probabilities nonlinearly into decision weights. We argue in Sect. 5 that this is a plausible reason for the difference between our findings in the domain of gains and those studies for farmers in developing countries that report finding an inverse S-shaped PWF.

⁴ More generally, on the limitations of proposed functional forms of the PWF for capturing experimentally observed behaviour, see Neilson and Stowe (2002).

We also contribute to the fairly small literature on probability weighting in the domain of losses (Tversky and Kahneman 1992; Abdellaoui 2000; Abdellaoui et al. 2005; Etchart-Vincent 2004; Abdellaoui et al. 2011). We make a small theoretical contribution by spelling out common consequence conditions for probability weighting in the domain of losses, make a methodological contribution by providing real incentives in the domain of losses without practising deception (cf. Etchart-Vincent and l'Haridon 2011), and expand the evidence base on probability weighting in the domain of losses by investigating it for a nonstandard subject pool.

The paper proceeds as follows. In Sect. 2, we derive our hypotheses in the context of cumulative prospect theory and develop common consequence conditions for testing these hypotheses. Section 3 shows how we implemented these common consequence conditions in our experimental design, and contains details of auxiliary data collection as well as the study area, sampling, and other fieldwork implementation. Section 4 presents descriptive statistics, a balancing test across game conditions, and the univariate and multivariate analyses for testing our hypotheses. Section 5 discusses our main results in terms of the theoretical expectations and the related empirical literature, and concludes.

2 Theory and hypotheses

In this section, we derive from theory the hypotheses to be tested in our experiments, as well as the vehicle for hypothesis testing, so-called common consequence ladders. Our theoretical framework is cumulative prospect theory (CPT) (Tversky and Kahneman 1992), one of the most influential theories of decision-making under uncertainty, which allows precise tests of probability weighting separately in the domains of gains and losses. The tests we derive from CPT consist of a series of choices between paired lotteries, each pair containing a relatively safe lottery and a relatively risky lottery. Each pair in the series is related to every other one through a single manipulation: an identical shift of probability mass between two outcomes, both in the safe lottery and in the risky lottery, a so-called common consequence shift (Wu and Gonzalez 1998). When a common consequence shift leads to a preference reversal (in one pair of lotteries the safe lottery is preferred, in another pair the risky lottery), then evidence has been obtained about the curvature of the probability weighting function (PWF) in a precisely defined interval of probabilities. In what follows, we first present decision weights in cumulative prospect theory, which make use of weighted cumulative probabilities. We next present S-shaped and inverse S-shaped PWFs, as well as the psychological intuitions that underpin them, and the hypotheses for our experimental tests that these give rise to. In the last two subsections, we present common consequence effects, in the domains of gains and losses, respectively. The common consequence conditions for probability weighting obtained in these two subsections form the basis for our experimental design, which constructs common consequence ladders separately in the domains of gains and losses, in a between-subject design.

2.1 Decision weights in cumulative prospect theory

In CPT, individuals maximise a strictly increasing value function $v : X \rightarrow \mathbb{R}$, where X is a set of monetary outcomes, with the neutral outcome denoted as 0, so that gains relative to the neutral outcome are denoted as positive numbers and losses as negative ones. The outcomes x_i of uncertain prospect f are arranged in increasing order, $x_{-m}, \dots, x_0, \dots, x_n$, with $-m$ to -1 indexing losses, 0 the neutral outcome, and 1 to n indexing gains. All positive outcomes, encapsulated by f^+ , are multiplied by decision weights $\pi^+(f^+) = (\pi_0^+, \dots, \pi_n^+)$; all negative outcomes in f^- by decision weights $\pi^-(f^-) = (\pi_{-m}^-, \dots, \pi_0^-)$. π_0^+ and π_0^- are in effect redundant, since $v(x_0) = v(0) = 0$.

The value function is additive in gains and losses, so that:

$$V(f) = V(f^+) + V(f^-) = \sum_{i=0}^n \pi_i^+ v(x_i) + \sum_{i=-m}^0 \pi_i^- v(x_i). \tag{1}$$

For risky prospects given by probability distribution p_i , decision weights are defined by:

$$\pi_n^+ \equiv w^+(p_n), \tag{2}$$

$$\pi_{-m}^- \equiv w^-(p_{-m}), \tag{3}$$

$$\pi_i^+ \equiv w^+(p_i + \dots + p_n) - w^+(p_{i+1} + \dots + p_n), \quad 0 \leq i \leq n - 1, \tag{4}$$

$$\pi_i^- \equiv w^-(p_{-m} + \dots + p_i) - w^-(p_{-m} + \dots + p_{i-1}), \quad 1 - m \leq i \leq 0. \tag{5}$$

Our interest is in the PWFs w^+ and w^- , defined, respectively, for gains and losses.⁵ They convert probabilities $0 \leq p_i \leq 1$ into weighted probabilities $w^+(p_i)$ and $w^-(p_i)$ through strictly increasing functions that satisfy $w^+(0) = w^-(0) = 0$ and $w^+(1) = w^-(1) = 1$.

Expressions 2–5 describe the transformation of weighted probabilities into decision weights. The decision weights on the best gain (expression 2) and worst loss (expression 3) are equal to the weighted probabilities associated with these outcomes. The decision weight on other positive outcomes (expression 4) is equal to the weighted probability that an outcome is at least as high as x_i , $w^+(p_i + \dots + p_n)$, minus the weighted probability that the outcome is strictly better, $w^+(p_{i+1} + \dots + p_n)$. The decision weight on negative outcomes other than the worst outcome (expression 5) is defined analogously as the difference between the weighted probability that an outcome is at least as bad as x_i , $w^-(p_{-m} + \dots + p_i)$ and the weighted probability that it is strictly worse, $w^-(p_{-m} + \dots + p_{i-1})$.

In Sects. 2.3 and 2.4, we specify decision weights π_n^+ , π_i^+ , π_{-m}^- , and π_i^- for three-outcome lotteries, and show how tests for nonlinear probability weighting may be derived from them. However, first, we wish to devote some space to S-shaped and

⁵ As will become clear later, tests for probability weighting using common consequence effects control for features of the value function, which, therefore, do not concern us here.

inverse S-shaped PWFs, and the psychological intuitions underpinning these PWFs, since these provide the hypotheses for our experimental tests.

2.2 S-shaped and inverse s-shaped probability weighting functions

The probability weighting postulated in cumulative prospect theory (Tversky and Kahneman 1992), as it was in original prospect theory (Kahneman and Tversky 1979), is that according to an inverse S-shaped PWF (Fig. 1). The psychological intuition invoked by Kahneman and Tversky is that of “diminishing sensitivity” relative to the reference points of $p^* = 0$ and $p^* = 1$, where the superscript * is used to indicate a reference probability. The idea is that changes near these reference probabilities are felt more acutely than changes further away from them: sensitivity to such changes diminishes as they take place further away from these reference points. Therefore, for example, a farmer considering the prospect of an investment in fertiliser would mind more (the value of the prospect would be more affected) if the probability of a loss changes from 5 to 10% than if it changed from 25 to 30%. Likewise, that farmer would mind more if the probability of a gain resulting from the investment dropped from 95 to 90% than if it dropped from 65 to 60%, if she weighted probabilities according to an inverse S-shaped PWF.

The bulk of the empirical evidence obtained in Western labs is consistent with an inverse S-shaped PWF, but there are exceptions.⁶ By contrast, the limited evidence available for rural areas of developing countries is sometimes consistent with an inverse S-shaped PWF (Tanaka et al. 2010; Liu 2013; Vieider et al. 2018), and sometimes consistent with an S-shaped PWF (Humphrey and Verschoor 2004a, b).⁷ In S-shaped PWFs (Fig. 2), changes in a middle interval of probabilities are felt more acutely than changes near the end points. Compare for example in Fig. 2 the increase in $w(p)$ of size a when p increases from 0.25 to 0.5 to the increase in $w(p)$ of size b when p increases from 0.75 to 1.0. Therefore, a farmer considering an investment in fertiliser would weight the probability increase of a gain from 50 to 60% more than the increase from 80 to 90%. Similarly, the probability of a loss decreasing from 50 to 40% would matter more than a decrease from 20 to 10% if probabilities are weighted according to an S-shaped PWF.

An S-shaped PWF is consistent with a reference probability somewhere in between the extremes of its domain. The psychological intuition underpinning that of a reference probability somewhere in the middle of its domain could be an awareness of the pervasive nature of risk. For instance, if the reference probability is informed by the historical relative frequency of outcomes, then in particularly hazardous environments, in which outcomes seldom occur with probabilities near 0 or 100%, a reference probability (or several reference probabilities) in between these two extremes is plausible.

⁶ See footnote 2.

⁷ As mentioned in the introduction, Tanaka et al. (2010), Liu (2013), and Vieider et al. (2018) use functional forms for the PWF that are limited in their tracking ability, whereas the probabilities in Humphrey and Verschoor (2004a, 2004b) are fairly crude (0.25, 0.5, and 0.75). Here, we advance the literature through more-refined tracking of the PWF both in the domain of gains and in the domain of losses.

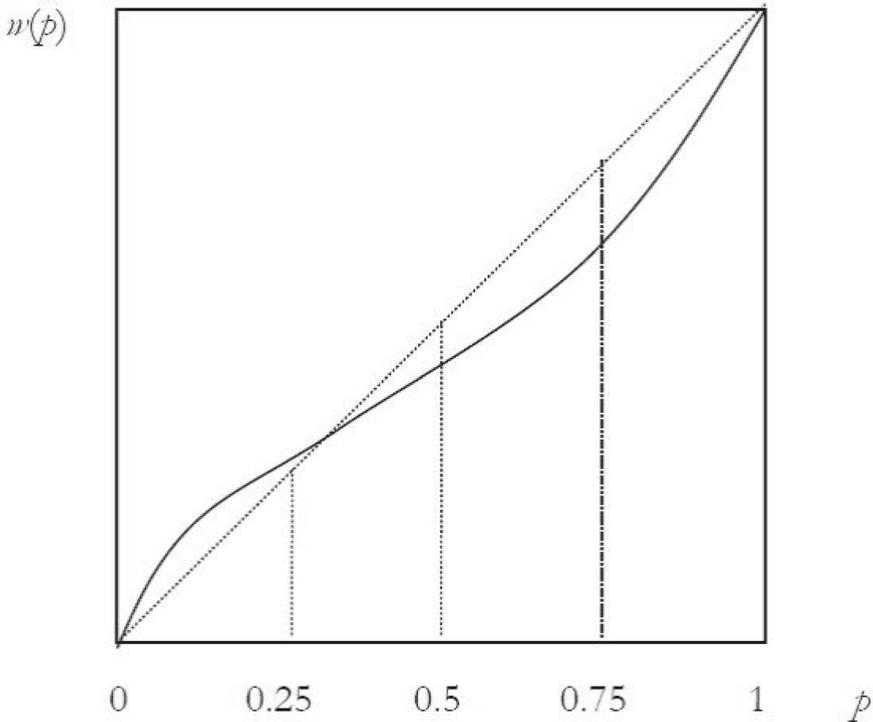


Fig. 1 An inverse S-shaped PWF

Diminishing sensitivity relative to reference probability p^* , $0 < p^* < 1$, would give rise to an S-shaped PWF: convex below p^* and concave above p^* , giving rise to a relatively steep curve in the vicinity of p^* and relative flatness near the end points of $p = 0$ and $p = 1$. In this study, we are interested in whether $0 < p^* < 1$ in a sample of Ugandan farmers. The reason we think this is plausible is that certainty ($p = 0$, $p = 1$) is at best an abstract concept in the livelihood decisions of these farmers. As detailed in Sect. 4.1, sharp negative income shocks due to harvest failure, droughts, and floods are very common in the study area. If certainty of outcomes is not remotely realistic, then a reference probability p^* other than $p^* = 0$ and $p^* = 1$, so $0 < p^* < 1$ might be plausible.

We summarise these considerations by spelling out the following four hypotheses, which we test in the paper. An inverse S-shaped PWF gives rise to the first two hypotheses.

Hypothesis 1 *The PWF is concave for probabilities $[0, p^y)$.*

Hypothesis 2 *The PWF is convex for probabilities $(p^y, 1]$.*

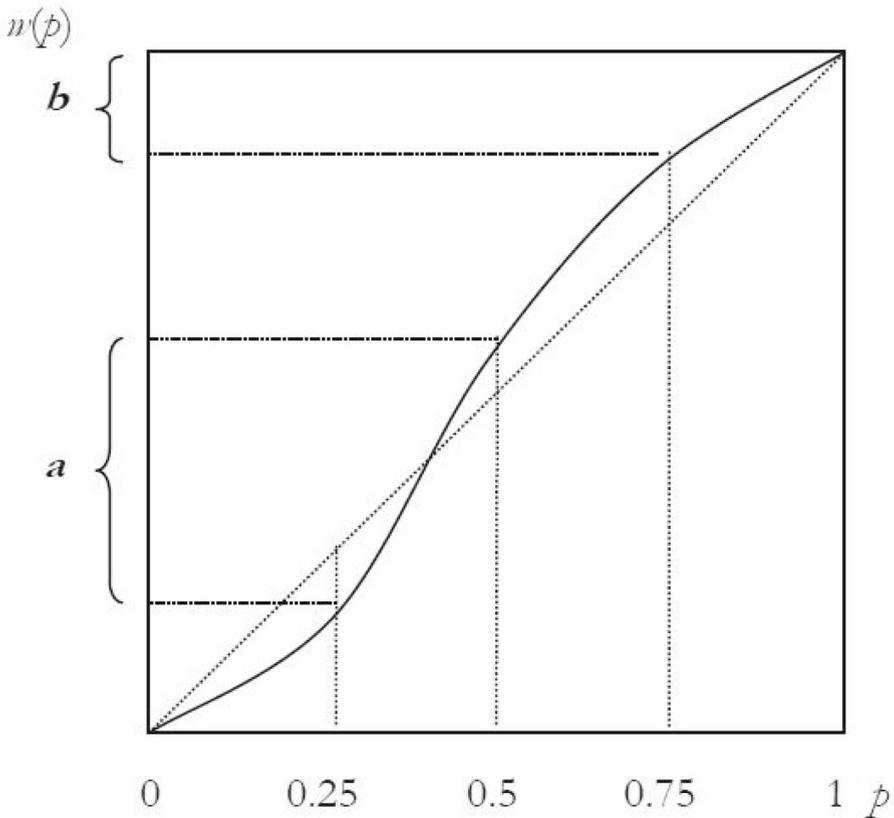


Fig. 2 An S-shaped PWF

Note that theory provides no guidance as to the magnitude of inflection point p^y ; the bulk of the empirical evidence suggests that $p^y \approx 1/3$.⁸

By contrast, an S-shaped PWF implies the next two hypotheses.

Hypothesis 3 *The PWF is concave for probabilities $(p^y, 1]$.*

Hypothesis 4 *The PWF is convex for probabilities $[0, p^y)$.*

In these two hypotheses, inflection point p^y is also reference probability p^* , with $0 < p^* < 1$. Again, theory provides no guidance as to its magnitude, so empirical tests should allow for it to lie anywhere in a considerable interval if they are to detect it. We made use of prior evidence on the likelihood of profits and losses, along with some desirable design features, for the specification of the intervals (one for gains and one for losses) within which to search for reference probabilities (see Sect. 3.2).

⁸ Among the reviews of studies of probability weighting, both Prelec (1998) and Starmer (2000) point this out.

All four hypotheses are evaluated relative to the null hypothesis of no probability weighting, which implies that the PWF is a straight line connecting $(p, w(p)) = (0, 0)$ and $(p, w(p)) = (1, 1)$. The null hypothesis follows from the independence axiom of expected utility theory. We next show how common consequence effects are useful for testing these hypotheses, first for gains and then for losses.

2.3 Common consequence effects in the domain of gains

For deriving conditions for the curvature of the PWF (convex or concave), we make use of common consequence effects. We limit the discussion to three-outcome lotteries, since these feature in our experimental design. The lotteries we consider are of the form $(p_0, x_0; p_1, x_1; p_2, x_2)$, where x_0 denotes the neutral outcome and $x_2 > x_1 > x_0$. For ease of comparison with Wu and Gonzalez (1998), we follow their notation and rewrite $x_2 \equiv x, x_1 \equiv y, x_0 \equiv z; p_2 \equiv p, p_1 \equiv q$, and $p_0 \equiv r$. The three-outcome lotteries we consider may thus be denoted $(p, x; q, y)$ in which the best outcome x occurs with probability p , the intermediate outcome y with probability q , and the worst (and neutral) outcome z with probability $r = 1 - p - q$.

In the domain of gains, the value function of CPT for the three-outcome prospects we consider may be written as:

$$V(f^+) = \pi_x^+ v(x) + \pi_y^+ v(y) + \pi_z^+ v(z). \tag{6}$$

From expressions 2 and 4, it follows that $\pi_x^+ = w^+(p)$ and $\pi_y^+ = w^+(p + q) - w^+(p)$. Since $v(z) = 0$, the value function may therefore be rewritten as:

$$v = w^+(p)v(x) + (w^+(p + q) - w^+(p))v(y). \tag{7}$$

This value function is identical to the one obtained in Wu and Gonzalez (1998), who derive their common consequence conditions in the context of rank-dependent expected utility (RDEU) (Quiggin 1982; Yaari 1987; Segal 1989). In the domain of gains, RDEU is indistinguishable from CPT. However, in the domain of losses, RDEU does not provide separate predictions: it simply ranks outcomes from worst to best and does not consider them relative to a neutral outcome. Following a slightly different approach, we therefore obtain in this subsection identical common consequence conditions for gains to those obtained by Wu and Gonzalez (1998), and derive novel common consequence conditions in the next subsection, for losses.

Consider a choice between safe lottery S (S for safe), characterised by (p', q') and risky lottery R (R for risky), characterised by (p, q) . Compared to S , R has a higher probability of the best outcome ($p > p'$) and a higher probability of the worst outcome ($1 - p - q > 1 - p' - q'$). Now, consider a choice between a second pair of lotteries S_ϵ and R_ϵ , which are derived from S and R , respectively, by shifting probability mass ϵ from the worst to the intermediate outcome.⁹ S_ϵ is thus

⁹ In Machina's unit probability triangle, this amounts to an identical horizontal translation of S and R ; see Wu and Gonzalez (1998, p. 123). The common consequence effects we consider in this subsection are therefore known as horizontal common consequence effects.

characterised by $(p', q' + \epsilon)$ and R_ϵ by $(p, q + \epsilon)$. A series of binary lottery choices constructed along these lines is known as a common consequence ladder: each lottery choice in the ladder is connected to every other one through a shift of probability mass between the same two outcomes, both in the safe and in the risky lottery. Importantly, outcomes x , y , and z are held constant, so that features of the value function are controlled for in common consequence conditions for the curvature of the PWF.

Nonlinear probability weighting may be inferred from preference reversals upon a common consequence shift. Consider the preference reversal $R \prec S$, while $R_\epsilon \succ S_\epsilon$, so the safe lottery is preferred before and the risky lottery after common consequence shift ϵ , $\epsilon > 0$. Such a preference reversal is an example of a common consequence effect. The common consequence effect considered here implies that the value of the risky prospect has risen more than the value of the safe prospect as a result of the common consequence shift, or $\Delta v = v_\epsilon - v > \Delta v' = v'_\epsilon - v'$. For spelling out implications for the curvature of the PWF, we obtain expressions for Δv and $\Delta v'$, using Eq. 7. The value of the risky lottery R , so before the common consequence shift, may be rewritten as follows:

$$v = w^+(p)[v(x) - v(y)] + w^+(p + q)v(y). \quad (8)$$

Likewise, the value of the risky lottery R_ϵ , so after the common consequence shift, is equal to $v_\epsilon = w(p)v(x) + (w(p + q + \epsilon) - w(p))v(y)$, which may be rewritten as:

$$v_\epsilon = w^+(p)[v(x) - v(y)] + w^+(p + q + \epsilon)v(y). \quad (9)$$

The increase in the value of the risky lottery as a result of the common consequence shift may thus be expressed as:

$$\Delta v = v_\epsilon - v = (w^+(p + q + \epsilon) - w^+(p + q))v(y). \quad (10)$$

Taking similar steps, the increase in the value of the safe lottery brought about by the common consequence shift is equal to:

$$\Delta v' = v'_\epsilon - v' = (w^+(p' + q' + \epsilon) - w^+(p' + q'))v(y). \quad (11)$$

The preference reversal $R \prec S$ while $R_\epsilon \succ S_\epsilon$ implies that $\Delta v > \Delta v'$, which by Eqs. 10 and 11 is equivalent to:

$$w^+(p + q + \epsilon) - w^+(p + q) > w^+(p' + q' + \epsilon) - w^+(p' + q'). \quad (12)$$

This inequality may be used to pronounce on the curvature of the PWF for specific intervals of its domain. Consider the interval $[p + q, p' + q' + \epsilon]$, in which $p + q < p' + q'$ (which we ensure in the experimental design). Strict concavity of $w^+(\cdot)$ in interval $[p + q, p' + q' + \epsilon]$ ensures that inequality 12 holds, since, by the definition of concavity, the PWF is steeper for the interval of probabilities $[p + q, p + q + \epsilon]$ than for the interval $[p' + q', p' + q' + \epsilon]$, which lies to the right of the first-mentioned interval. Strict concavity of $w^+(\cdot)$ in the interval $[p + q, p' + q' + \epsilon]$ is thus a sufficient condition for inequality 12 to hold. For it to be a necessary

condition, more structure needs to be imposed.¹⁰ Prudently, one should, therefore, limit pronouncements on the curvature of the PWF in the light of preference reversals $R \prec S$, while $R_\epsilon \succ S_\epsilon$ to stating that evidence has been obtained that the PWF is steeper in the interval $[p + q, p + q + \epsilon]$ than in the interval $[p' + q', p' + q' + \epsilon]$, while remaining agnostic about concavity in the entire interval of its domain $[p + q, p' + q' + \epsilon]$.

Mirroring the steps taken to obtain inequality 12 by reversing the inequality signs in these steps, we may state that the preference reversal $R \succ S$, while $R_\epsilon \prec S_\epsilon$ implies the following inequality:

$$w^+(p + q + \epsilon) - w^+(p + q) < w^+(p' + q' + \epsilon) - w^+(p' + q'), \tag{13}$$

or that the PWF is steeper in the interval $[p' + q', p' + q' + \epsilon]$ than in the interval $[p + q, p + q + \epsilon]$.

2.4 Common consequence effects in the domain of losses

We next derive common consequence conditions for the curvature of the PWF in the domain of losses, in the context of CPT. We consider three-outcome lotteries of the form $(p, x; q, y; r, z)$ in which the worst outcome x occurs with probability p , the intermediate outcome y with probability q , and the neutral outcome z with probability $r = 1 - p - q$. So now, $x < y < z$ and $v(z) = 0$.

In the domain of losses, the value function of such three-outcome lotteries may be written as:

$$V(f^-) = \pi_x^- v(x) + \pi_y^- v(y) + \pi_z^- v(z). \tag{14}$$

From expressions 3 and 5, it follows that $\pi_x^- = w^-(p)$ and $\pi_y^- = w^-(p + q) - w^-(p)$. Since $v(z) = 0$:

$$v = w^-(p)v(x) + (w^-(p + q) - w^-(p))v(y). \tag{15}$$

The problem of obtaining common consequence conditions is thus symmetrical in its starting point to that for gains. However, since $v(x) < 0$ and $v(y) < 0$, some of the key inequalities reverse. Moreover, the common consequence shifts we consider, from the worst to the intermediate outcome, give rise to changes in the value of prospects that are not symmetrical to such changes in the domain of gains, as will be seen below.

Consider a choice between safe lottery S , $(p', x; q', y)$ and risky lottery R , $(p, x; q, y)$, in which $p' < p$ and $1 - p' - q' < 1 - p - q$, i.e., the probability of the worst outcome occurring and that of the best (neutral) outcome occurring are both lower in S than in R . By construction, $q' > q$.

Our interest is in the curvature of w^- , which we infer from the absence or presence of preference reversals between various pairs of lotteries that relate to each

¹⁰ Wu and Gonzalez (1998, pp. 125ff.) show that for small ϵ , continuous and twice differentiable $w(\cdot)$, a local condition $w''(p + q) < 0$ is approached as $p' + q' \rightarrow p + q$.

other through common consequence shifts in the form of a shift of probability mass from the worst to the intermediate outcome. S_ϵ is thus characterised by $(p' - \epsilon, q' + \epsilon)$ and R_ϵ by $(p - \epsilon, q + \epsilon)$.

Consider the preference reversal $R \prec S$ while $R_\epsilon \succ S_\epsilon$ after the common consequence shift $\epsilon, \epsilon > 0$.¹¹ As before, this implies that $\Delta v = v_\epsilon - v > \Delta v' = v'_\epsilon - v'$. For the risky lottery, the common consequence shift implies that:

$$v = w^-(p)v(x) + (w^-(p + q) - w^-(p))v(y) = w^-(p)[v(x) - v(y)] + w^-(p + q)v(y); \tag{16}$$

after the common consequence shift becomes:

$$\begin{aligned} v_\epsilon &= w^-(p - \epsilon)v(x) + (w^-(p + q) - w^-(p - \epsilon))v(y) \\ &= w^-(p - \epsilon)[v(x) - v(y)] + w^-(p + q)v(y) \end{aligned} \tag{17}$$

(since $p - \epsilon + q + \epsilon = p + q$). The increase in value of the risky lottery as a result of the common consequence shift may thus be written as:

$$\Delta v = v_\epsilon - v = [w^-(p - \epsilon) - w^-(p)][v(x) - v(y)] > 0, \tag{18}$$

which is strictly positive, since $w^-(p - \epsilon) < w^-(p)$ and $v(x) < v(y)$.

Analogously, for safe lotteries S and S_ϵ :

$$\Delta v' = v'_\epsilon - v' = [w^-(p' - \epsilon) - w^-(p')][v(x) - v(y)] > 0. \tag{19}$$

Preference reversal $R \prec S$ while $R_\epsilon \succ S_\epsilon$ implies:

$$\Delta v > \Delta v' \Leftrightarrow [w^-(p - \epsilon) - w^-(p)][v(x) - v(y)] > [w^-(p' - \epsilon) - w^-(p')][v(x) - v(y)]. \tag{20}$$

Dividing both sides by $[v(x) - v(y)] < 0$ reverses the inequality and gives the following common consequence condition:

$$R \prec S, R_\epsilon \succ S_\epsilon \Rightarrow \Delta v > \Delta v' \Leftrightarrow w^-(p - \epsilon) - w^-(p) < w^-(p' - \epsilon) - w^-(p'). \tag{21}$$

Since $w^-(p - \epsilon) < w^-(p)$ and $w^-(p' - \epsilon) < w^-(p')$, both sides of the inequality are negative. It may therefore be more natural to divide both sides by minus one and write the common consequence condition as follows:

$$R \prec S, R_\epsilon \succ S_\epsilon \Rightarrow \Delta v > \Delta v' \Leftrightarrow w^-(p) - w^-(p - \epsilon) > w^-(p') - w^-(p' - \epsilon). \tag{22}$$

In words, the preference reversal implies that the PWF for losses is steeper in the interval of its domain $[p - \epsilon, p]$ than in $[p' - \epsilon, p']$. Since $p' < p$, strict convexity in

¹¹ Such a common consequence effect occurs in response to a vertical translation of lotteries R and S in Machina's unit probability triangle, and is therefore known as a vertical common consequence effect. See Wu and Gonzalez (1998, p. 123) for an example in the domain of gains.

the interval of its domain $[p' - \epsilon, p]$ is a sufficient condition for $w^-(p) - w^-(p - \epsilon) < w^-(p') - w^-(p' - \epsilon)$ but not a necessary condition (cf. the discussion on the corresponding concavity condition in Sect. 2.3), so we limit our empirical conclusions to the relative steepness of the PWF for interval $[p - \epsilon, p]$ compared to that of $[p' - \epsilon, p']$.

Working through the same logic for the opposite preference reversal yields the mirroring common consequence condition:

$$R \succ S, R_\epsilon \prec S_\epsilon \Rightarrow \Delta v < \Delta v' \Leftrightarrow w^-(p) - w^-(p - \epsilon) < w^-(p') - w^-(p' - \epsilon), \quad (23)$$

so that the PWF for losses is steeper in the interval $[p' - \epsilon, p']$ than in the interval $[p - \epsilon, p]$.

We will next show how our experimental design makes use of the common consequence conditions obtained in Sects. 2.3 and 2.4 to test the hypotheses presented in Sect. 2.2.

3 Experimental design, survey, and fieldwork implementation

In this section, we describe our data collection instruments. We first show how we implemented common consequence ladders in a sample with low levels of literacy, present the lottery choices we designed, and provide a rationale for the intervals of probabilities and magnitudes of the common consequence shifts we focused on. Next, we describe the steps we took to establish a neutral outcome in subjects' minds in the weeks leading up to the experiment, so that gains and losses were meaningful concepts to them when they chose between lotteries. We then present the other key elements of our experimental design: random assignment to either the gains or losses version of the common consequence ladders, and design choices as to the order of the experimental tasks, a simple control for risk aversion, and the random selection of one task for payment. The experiment was complemented with a survey, which we outline in the last subsection, along with details of sample selection and fieldwork implementation.

3.1 Implementing common consequence ladders

Literacy and the ability to translate visual displays on a computer screen into concrete realisations cannot be guaranteed in the study area, so we needed a device for implementing pairwise lottery choices that does not rely on written instructions, and is concrete and simple. Based on extensive piloting, we settled on the following device. Each lottery was represented by twenty coloured counters. So, for example, lottery 6a in the domain of gains (see Table 1) has a 0.4 chance of the neutral outcome, a 0.6 chance of the intermediate outcome, and a 0 chance of the best outcome. This lottery was represented by 8 lilac counters ($8/20 = 0.4$) and 12 light blue counters. These were vertically set out on a table in single file, so that the 8 lilac counters were neatly arranged at the bottom and the 12 light blue counters on top of those. Underneath that column of counters, 6a was clearly written on a large

Table 1 Common consequence ladder (gains)

Rung	Order	Risky or safe	Lilac @ 8000 (z)	Light blue @ 10,000 (y)	White @ 13,000 (x)
I	7a	S	9 ($r' = 0.45$)	11 ($q' = 0.55$)	0 ($p' = 0$)
	7b	R	13 ($r = 0.65$)	0 ($q = 0$)	7 ($p = 0.35$)
II	6a	S	8 (0.4)	12 (0.6)	0 (0)
	6b	R	12 (0.6)	1 (0.05)	7 (0.35)
III	10a	S	7 (0.35)	13 (0.65)	0 (0)
	10b	R	11 (0.55)	2 (0.1)	7 (0.35)
IV	5a	S	6 (0.3)	14 (0.7)	0 (0)
	5b	R	10 (0.5)	3 (0.15)	7 (0.35)
V	2a	S	5 (0.25)	15 (0.75)	0 (0)
	2b	R	9 (0.45)	4 (0.2)	7 (0.35)
VI	3a	S	4 (0.2)	16 (0.8)	0 (0)
	3b	R	8 (0.4)	5 (0.25)	7 (0.35)
VII	4a	S	3 (0.15)	17 (0.85)	0 (0)
	4b	R	7 (0.35)	6 (0.3)	7 (0.35)
VIII	9a	S	2 (0.1)	18 (0.9)	0 (0)
	9b	R	6 (0.3)	7 (0.35)	7 (0.35)
IX	8a	S	1 (0.05)	19 (0.95)	0 (0)
	8b	R	5 (0.25)	8 (0.4)	7 (0.35)
X	1a	S	0 (0)	20 (1)	0 (0)
	1b	R	4 (0.2)	9 (0.45)	7 (0.35)

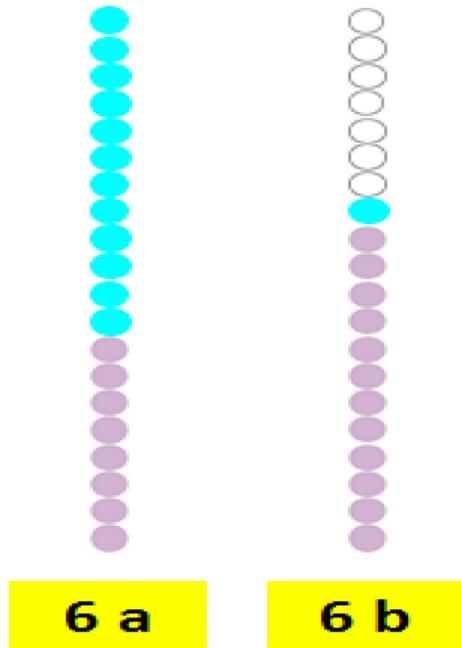
The table presents the 10 pairwise lottery choices in the gains condition of the experiment. The number of counters of each colour determines the probability of x (white), y (light blue), and z (lilac); probabilities of outcomes are in parentheses. The first column indicates the rung on the common consequence ladder each pair of lotteries represents; the second column the order in which they were presented (the reverse order for a randomly selected half of subjects)

post-it note stuck to the table. About 40 cm next to the column of counters representing lottery 6a was lottery 6b set out in similar fashion (12 lilac, 1 light blue, and 7 white counters). Figure 3 illustrates these two lotteries.

Before they took their decisions, a demonstration was given to subjects of what would happen once they chose one of the two lotteries. All counters representing that lottery would be put in a bag, shuffled thoroughly, and one would be picked out by a volunteer looking the other way. The value of the colour of the counter would be paid out to subjects who had chosen that lottery. We demonstrated this both for the relatively safe (6a) and for the relatively risky lottery (6b).

We ensured that all paired lotteries were spatially sufficiently apart, so that each lottery choice problem would be considered in isolation. For this purpose, we used five large tables arranged in the middle of the experiment room. On one side were set out lottery pairs 1–5, on the other 6–10, with plenty of space left in the middle between the lottery pairs on each side of the table. This arrangement not only

Fig. 3 An example of a lottery choice problem. Notes: this represents choice problem 6 in the gains condition of the experiment. Subjects were shown that the counters of the lottery they preferred would be put in a bag, from which one counter would be randomly selected. The value of the outcome that counter represents would be their payment. In the gain condition, lilac counters represented 8000 Ugandan shillings, light blue counters 10,000 shillings, and white counters 13,000 shillings



encouraged considering each choice in isolation, but also greatly facilitated reversing the order in which lottery choices were presented (see Sect. 3.4).

A control question designed to capture understanding of the essence of the device revealed that subject comprehension was very good: 96% answered it correctly.¹²

3.2 Common consequence ladders implemented

Table 1 presents the common consequence ladder implemented in the domain of gains. The first pair of lotteries (on rung I) constitutes the least attractive choice, the next pair (rung II) the least attractive choice but one, and so forth, until the most attractive choice in terms of expected value of the lotteries (rung X). Each pair consists of a relatively safe and a relatively risky lottery, which subjects face in the form of the number of counters of up to three colours presented in the last three columns. The probability of each outcome is indicated in parentheses.

These ten lottery choice problems were determined as follows. The comparison of choices between rungs allows pronouncing on the relative steepness of the PWF for specific intervals of probabilities. For the horizontal common consequence effects we consider, comparisons are between intervals $[p + q, p + q + \epsilon]$ and $[p' +$

¹² The control question asked for the gains version of the experiment was: “We just want to check your understanding of the task. Can you please tell me, of the two lotteries here in front of you, which one offers the higher chance of leaving with exactly 10,000 shillings?” For the losses version of the experiment, a very similar control question was asked. See the experimental instructions in the online appendix.

$q', p' + q' + \epsilon]$ with $p' + q' > p + q$ needing to be ensured in the design (see Sect. 2.3).

Consider for example a preference reversal between rung I and rung II. $p + q = 0.35$ refers to the probability of either the white or light blue counter being selected in I.R, $p' + q' = 0.55$ to the probability of that outcome in I.S, and common consequence shift ϵ is added to the analogous probabilities in I.I.S and I.I.R. A preference reversal therefore indicates the relative steepness of the PWF in probability intervals $[0.35, 0.4]$ compared to $[0.55, 0.6]$. Following the same logic, comparing rungs V and VI pronounces on the relative steepness of the PWF in intervals $[0.55, 0.6]$ and $[0.75, 0.8]$, rungs IX and X intervals $[0.75, 0.8]$ and $[0.95, 1.0]$, and so forth, for a total number of possible comparisons of $10 * 9/2 = 45$ across each combination of rungs.

Such tracking of the PWF for gains is thus limited to part of its domain $[0.35, 1.0]$, which we decided for the sake of realism, avoiding cognitive overload and prioritising more-refined measurement to measurement over the entire domain

Table 2 Common consequence ladder (losses)

Rung	Order	Risky or safe	Lilac @ 3000 (x)	Light blue @ 5000 (y)	White @ 8000 (z)
I	3a	S	14 ($p' = 0.7$)	6 ($q' = 0.3$)	0 ($r' = 0$)
	3b	R	16 ($p = 0.8$)	0 ($q = 0$)	4 ($r = 0.2$)
II	1a	S	13 (0.65)	7 (0.35)	0 (0)
	1b	R	15 (0.75)	1 (0.05)	4 (0.2)
III	6a	S	12 (0.6)	8 (0.4)	0 (0)
	6b	R	14 (0.7)	2 (0.1)	4 (0.2)
IV	9a	S	10 (0.5)	10 (0.5)	0 (0)
	9b	R	12 (0.6)	4 (0.2)	4 (0.2)
V	4a	S	8 (0.4)	12 (0.6)	0 (0)
	4b	R	10 (0.5)	6 (0.3)	4 (0.2)
VI	10a	S	6 (0.3)	14 (0.7)	0 (0)
	10b	R	8 (0.4)	8 (0.4)	4 (0.2)
VII	5a	S	4 (0.2)	16 (0.8)	0 (0)
	5b	R	6 (0.3)	10 (0.5)	4 (0.2)
VIII	2a	S	2 (0.1)	18 (0.9)	0 (0)
	2b	R	4 (0.2)	12 (0.6)	4 (0.2)
IX	7a	S	1 (0.05)	19 (0.95)	0 (0)
	7b	R	3 (0.15)	13 (0.65)	4 (0.2)
X	8a	S	0 (0)	20 (1)	0 (0)
	8b	R	2 (0.1)	14 (0.7)	4 (0.2)

Notes. The table presents the 10 pairwise lottery choices in the losses condition of the experiment. The number of counters of each colour determines the probability of x (lilac), y (light blue), and z (white); probabilities of outcomes are in parentheses. The first column indicates the rung on the common consequence ladder each pair of lotteries represents; the second column the order in which they were presented (the reverse order for a randomly selected half of subjects)

[0, 1.0]. As to realism, farmers' investments in fertiliser, improved seeds, and other pertinent agricultural investment goods in Uganda are successful with a probability considerably higher than 0.35 (Verschoor et al. 2016, Table A1, p.147). Avoiding cognitive overload made us decide after piloting that no more than 10 rungs should be implemented and that each lottery should have no more than 20 counters. As to refined measurement, limiting tracking of the PWF to part of its domain [0.35, 1.0] meant that we could set $\epsilon = 0.05$ between each two adjacent rungs in the common consequence ladder for gains.

Table 2 presents the common consequence ladder implemented in the domain of losses. Pairs of lotteries are again presented in order from the one with the lowest expected values of S and R on rung I to the one with the highest expected value on rung X. For losses, comparisons across paired lotteries are relevant for the relative steepness of the PWF between probability intervals $[p' - \epsilon, p']$ and $[p - \epsilon, p]$ (see Sect. 2.4).

For example, in lottery I.R, $p = 0.8$ and in lottery I.S, $p' = 0.7$. When the common consequence shift of size $\epsilon = 0.05$ is considered between paired lotteries I and paired lotteries II, then $p' - \epsilon = 0.65$ and $p - \epsilon = 0.75$. A preference reversal from R to S or S to R between lotteries I and lotteries II thus indicates that the PWF is steeper in one of the two probability intervals [0.65, 0.7] and [0.75, 0.8] than it is in the other.

Following similar logic, comparing choices in V and VI pronounces on the relative steepness of the PWF in intervals [0.3, 0.4] and [0.4, 0.5], comparing choices in IX and X on [0, 0.05] and [0.1, 0.15], and so forth, again for a total possible number of comparisons of $10 * 9/2 = 45$.

The tracking of the PWF for losses is thus restricted to part of its domain [0, 0.8]. Considerations of realism, refinement where it matters most, and avoiding cognitive overload have again inspired this design choice. The last two mentioned are similar to those considerations for gains. As to realism, sharp negative income shocks are common in the study area (see Sect. 4.1), which suggests sizeable reference probabilities somewhere in the middle of the PWF's domain (cf. Sect. 2.2). To reduce the risk of failing to detect reference probabilities, we therefore cast our net somewhat wider than in the domain of gains. Moreover, whereas for gains, it seemed important to close in on certainty at one end of the PWF's domain (a 100% certain gain); for losses, it seemed important to close in on certainty at the other end (a zero% chance of a loss).

The two common consequence ladders presented here are valid to the extent that we have successfully established in subjects' minds that $z = 8,000$ Ugandan shillings are thought of as the neutral outcome. We next show how we attempted this.

3.3 Establishing the neutral outcome

Three weeks before a subject was scheduled to participate, we gave them a voucher. The voucher contained their name, address, and photo, as well as the figure of 8000

shillings prominently displayed. The following scripted message was read out to participants individually when the voucher was handed over:

On the day of the research workshop you'll be asked to take some decisions about this money. Depending on the decisions you will be asked to take, you could end up with more than 8000 or with less than 8000 shillings.

During the delivery of the experimental instructions 3 weeks later, the voucher was referred to when the outcome of 8000 shillings was introduced (see online Appendix A, p.8):

Remember, you have already been given a voucher worth 8000 shillings. Therefore, if a white counter is eventually drawn, then you do not earn any extra money, but keep your 8000 shillings.

The vouchers were thus intended to instil a sense of entitlement. Deception was avoided through emphasising that the money could be lost, depending on the decisions subjects were asked to take. To avoid disappointment, an unannounced show-up fee of 5000 shillings was paid after the resolution of the game. The amount of 5000 shillings was the maximum amount that could be lost in the losses condition. Minimum earnings were therefore 8000 shillings, or about twice the median daily wage in the study area. Crucially though, losses could actually be made out of the 8000 shillings mentioned on the voucher, exactly as per the experimental instructions that subjects responded to.¹³

3.4 Other elements of the experimental design

Other elements of the design were as follows. We implemented a simple investment game so as to be able to control for risk aversion, based on the Gneezy and Potters (1997) design. Subjects were endowed with 20 counters, each representing 400 shillings, so 8000 shillings in total. They chose to invest k counters, where $k \in \{0, 1, \dots, 20\}$ for facing the lottery $(0.5, 8000 - 400k; 0.5, 8000 + 800k)$. In other words, their investment was tripled if successful and lost in its entirety if it failed.¹⁴ The fate of their investment is determined by tossing a coin. We assume a power Constant Relative Risk Aversion (CRRA) utility function over experimental earnings x , which is defined as $U(x) = x^{1-r}/(1-r)$, where r is the coefficient of CRRA. As is conventional, we compute the CRRA coefficient for indifference between investing k and $k - 1$, on one hand, and k and $k + 1$ on the other, to find the CRRA coefficient range that corresponds with the observed behaviour of investing

¹³ When losses are implemented in experiments, it is customary to either work with hypothetical payoffs (e.g., Etchart-Vincent 2004; Abdellaoui et al. 2005), or to practise mild forms of deception when implementing real incentives (e.g., Yesuf and Bluffstone 2009). Both are done for ethical reasons, so as to ensure that subjects have no legitimate grounds for feeling deprived as a result of their participation in the experiment. The implementation of hypothetical losses has been compared in one study with that of real incentives, and little difference was found (Etchart-Vincent and l'Haridon 2011). Even so, we could not be certain that this would hold in our sample, so we implemented real losses from an initial endowment, while avoiding deception and the house-money effect (Thaler and Johnson 1990).

¹⁴ The rate of return on investment was calibrated during the pilot for inducing variation in behaviour.

k. The investment decision was always task 11, so after the 10 lottery choices of the main experiment had taken place.

The random-lottery incentive system was used, so that only one decision was implemented and each of the 11 lottery choices (the decision in the investment game is in effect also a lottery choice) had an equal chance of being played out for real. Subjects were told this before they took their decisions.

The order of the ten paired lottery choices was randomised, both in the gains condition and in the losses condition (see Tables 1 and 2). Moreover, whether a subject faced the sequence of lottery choices 1, 2, ..., 10 or that of 10, 9, ..., 1 was randomly determined.

Two experimental teams were used for delivering the instructions, who were purposively rotated across sessions and game conditions, so that we can control for experimenter effects. The experimental teams are experienced and were intensively trained for 2 weeks, in addition to their involvement in extensive piloting.

Subjects were randomly assigned to either the gains condition or the losses condition of the experiment in a between-subject design.

Experimental instructions were delivered in person, because literacy cannot be guaranteed in the study area. Decision-making was partially private: other subjects were not present but an enumerator recorded choices.

The experimental instructions were translated into Lugisu, the local language of the study area, and back-translated into English to check for inadvertent changes in meanings.

A number of instructions and reassurances were given at the outset to promote orderly, leisurely, and autonomous decision-making; see the experimental instructions in online Appendix A.

3.5 Study area, sample selection, survey, and fieldwork implementation

We selected five sub-counties from a rural area in eastern Uganda: Sironko District and Lower Bulambuli District, which together comprise the former Sironko District. Almost all working adults in the area (95%) are smallholder farmers. The typical farmer grows maize intercropped with beans alongside some cash crops (e.g., coffee) on land that does not exceed 2 acres. Using a multi-stage cluster sampling method, we selected 1803 farmers for a number of risky choice experiments including the ones reported on here. Of the individuals selected for participation, 370 were randomly assigned to the common consequence experiments, 184 of them to the gains condition, and 186 to the losses condition. For a detailed description of the study area and fieldwork implementation, see Verschoor et al. (2016).

Data collection followed the same basic pattern in each sub-county. Three weeks before their relevant experimental session, individuals were visited, reminded about their participation and given the voucher described in Sect. 3.3. In the week before 'game day', we visited all selected participants again to administer a household survey questionnaire, to collect data on basic socio-economic variables, agricultural practises, and the experience of income shocks. The experiment then took place at the end of the week in a central location (usually in a school on a nonschool day); for participants from remote villages, we organised transport.

4 Results

In this section, we present our main results. We first show summary statistics and ascertain whether condition assignment has given rise to inadvertent selection issues. Next, we analyse common consequence effects in the domains of gains and losses, respectively, and derive the shape of the PWFs (one for losses, one for gains) that organise the data. Finally, we present the results of multivariate analysis that searches for correlates of the particularly pronounced nonlinear probability weighting that the univariate analysis has pointed to.

4.1 Sample characteristics

We begin by describing our sample and assess differences on observables between game conditions. Table 3 presents summary statistics and a corresponding balancing test. About 30% of subjects are traditional farmers. We consulted

Table 3 Summary statistics and balancing test

Variable	Mean (standard deviation)			<i>t</i> statistic	<i>p</i> value
	Total (<i>N</i> =370)	Gains condition (<i>N</i> =184)	Losses condition (<i>N</i> =186)		
Traditional farmer (=1)	0.30 (0.46)	0.30 (0.46)	0.31 (0.46)	-0.12	0.90
Wealth index	0.06 (2.38)	-0.20 (1.88)	0.31 (2.77)	-2.06	0.04
Risk aversion (CRRA)	1.29 (2.01)	1.40 (2.26)	1.17 (1.74)	1.10	0.27
Female (=1)	0.52 (0.50)	0.55 (0.50)	0.48 (0.50)	1.31	0.20
Years of schooling	5.46 (3.59)	5.37 (3.51)	5.55 (3.68)	0.49	0.62
Age	40.85 (14.03)	40.50 (14.45)	41.20 (13.63)	0.48	0.63
Harvest failed in past 5 years (=1)	0.77 (0.42)	0.78 (0.41)	0.75 (0.43)	0.65	0.52
Experienced severe flooding in past 5 years (=1)	0.59 (0.49)	0.60 (0.49)	0.58 (0.50)	0.40	0.69
Experienced severe drought in past 5 years (=1)	0.84 (0.37)	0.85 (0.36)	0.83 (0.38)	0.49	0.62
Control question correct (=1)	0.96 (0.31)	0.99 (0.35)	0.93 (0.26)	1.87	0.06

Notes. The table presents summary statistics and a test of the null hypothesis that mean values are equal by game condition. In a logistic regression of condition assignment on all variables presented in the table, the null hypothesis of all coefficients being jointly equal to zero cannot be rejected ($\chi^2 = 10.8$; $p = 0.46$)

agricultural experts in the study area to define traditional farmers.¹⁵ Traditional farmers grow maize intercropped with beans, with a minimal reliance on bought inputs such as improved seeds, pesticides, and inorganic fertiliser; if they do buy inputs, then this would be limited to purchasing unimproved seeds on local markets. We therefore classified all subjects as traditional farmers who grow only maize and beans and do so without buying improved seeds, pesticides, or fertiliser. All other subjects either buy such inputs or grow cash crops that are more lucrative than maize and beans, but require higher cash outlays (e.g., tomatoes, cabbages, and onions) or both. According to the local experts, traditional farmers avoid initial cash outlays so as not to jeopardise their food security in periods of drought or excessive flooding. This is a recognised motive in development economics (Fafchamps 2003, pp.18ff.).¹⁶

The wealth index presented in the table is the first component of a principal component analysis based on a list of about 30 types of assets that our survey collected information on (this method of constructing a wealth index is due to Filmer and Pritchett (2001)). Risk aversion is measured using the constant relative risk aversion (CRRA) coefficient based on behaviour in the investment game presented in Sect. 3.4.

We asked several questions about income shocks. The way we asked these questions was by inquiring whether respondents had experienced a sharp drop in income in the past 5 years as a result of specified events. 77% had experienced a sharp drop in income resulting from harvest failure, 59% from severe flooding, and 84% from a severe drought. Farmers in the sample are clearly used to frequent sharp income shocks.

Most variables in Table 3 are not significantly different between game conditions, but two variables are. One is the wealth index. We think that this should be due to chance, as the random assignment to game condition was rigorously carried out. Game comprehension is the other variable that is significantly different between game conditions: 99% answered the control question correctly in the gains condition and 93% in the losses condition, which is significantly different at the 10% level. This could well be due to the differences between the game versions, with the losses instructions possibly slightly harder to understand. Reassuringly, the omnibus χ^2 -statistic of a test of all coefficients being equal to zero in a logistic regression of condition assignment on the variables presented in the table is insignificant ($\chi^2 = 10.8$, $p = .46$), suggesting that inadvertent selection is not a major concern.

Table 4 contains a summary of experimental behaviour. Choosing the risky lottery is somewhat more frequent than the safe one, especially for losses and, within the domain of losses, especially for the lower rungs. In the domain of losses, the gradual lowering of the proportion of risky choices as one moves up the ladder

¹⁵ These agricultural experts were interviewed individually and included the District Agricultural Officer, several agricultural extension workers, and several leaders of farmers' groups.

¹⁶ Traditional farmers are significantly less wealthy ($t = -7.47$, $p = 0.000$), more likely to be female ($t = 3.14$, $p = 0.002$), spent fewer years in school ($t = -5.87$, $p = 0.000$), and are older than others in the sample ($t = 2.64$, $p = 0.008$). The p values reported are based on two-sided tests.

Table 4 Summary of behaviour in experiments

Rung	I	II	III	IV	V	VI	VII	VIII	IX	X
% Risky, losses	63.6	61.2	57.6	58.2	54.9	61.4	58.7	55.4	48.4	52.7
% Switch from previous rung, losses		27.4	31.2	26.3	36.6	32.3	36.0	33.3	29.6	36.6
% Risky, gains	56.0	51.4	54.1	51.1	51.6	46.7	51.6	47.8	50.5	50.5
% Switch from previous rung, gains		28.8	24.5	34.2	33.2	26.6	32.1	26.6	28.8	32.6

to rung V implies a preponderance of *RS* switches (choosing risky on a lower rung followed by safe on the higher rung), which hints at the relative steepness of the PWF near $p = 0$; we formally test this below. Table 4 also shows that about a third of subjects choose different lotteries on consecutive rungs, which our tests for probability weighting in the next subsection focus on. In the domain of losses, 19.6% of subjects never deviate from their first choice; 13.0% always choose the risky lottery, and 6.5% always safe. In the domain of gains, the corresponding figures are 35.9, 19.0, and 16.8%. On average, experimental subjects switch about three times between consecutive rungs (recall that consecutive rungs were not sequentially presented, since the order in which they were presented to subjects was randomised). In the domain of losses, the mean number of switches between consecutive rungs is 2.9 (Std Dev 2.1), and for gains 2.7 (Std Dev 2.4).

4.2 Tracking the PWF

We next present the choice patterns that the common consequence shifts have given rise to, and show the shapes of the PWFs, one in the domain of losses and one in the domain of gains, that organise the data. Our tests of probability weighting used for tracking the PWF may be construed as tests of the null hypothesis of expected utility theory (EUT) plus noise arising from simple (Fechnerian) errors in decision-making (Hey and Orme 1994). For these tests to be valid, we need to maintain the assumption that erroneous *SR* switches are not more likely than erroneous *RS* switches, in any of the comparisons considered between rungs of the common consequence ladder.

Recall that there are two experimental conditions: about half of the subjects are in the gains condition, and the other half are in the losses condition. Each subject makes ten risky choice decisions. In each of the ten decision-making problems, they can choose either the safe (*S*) or the risky (*R*) lottery. As is standard in tests for common consequence effects, probability weighting is tested for by comparing decisions in paired problems. There are 45 possible comparisons ($10 * 9/2$), which all allow tests of probability weighting for specific ranges of probabilities (e.g., whether the probability weighting function is steeper for 0.8–1.0 than it is for 0.4–0.6).

Therefore, we test for probability weighting by focusing on paired problems. Let us consider two decision-making problems, problem 1 and problem 2. An individual decides whether they want S or R in problem 1, and next whether they want S or R in problem 2. There are therefore four possible choice patterns: SS , RR , SR , and RS .

The focus of the analysis (again standard in tests of common consequence effects) is on the proportion of subjects that chooses each of these four possibilities. The null hypothesis is: no probability weighting. In Sect. 2, we showed (in line with standard theory) that this means that if somebody chooses S in one problem, then they should also choose S in the second problem (and vice versa). Likewise, if they choose R in one problem, then they should also choose R in the second problem (and vice versa). The null hypothesis of no probability weighting and no errors requires all choices to be SS or RR .

For testing this, we allow subjects to make mistakes. Therefore, the null hypothesis becomes: apart from mistakes, all choices should be SS or RR . We then make the assumption that mistakes in one direction (SR) are equally likely as mistakes in another direction (RS). For testing for probability weighting, there are several possibilities.

First, we could assume that the lowest of SR and RS is entirely due to mistakes. Using a one-sample binomial test, we can then test whether the higher of RS and SR significantly differs from the lower, the lower being the hypothesized value due to mistakes. This is defensible, but there is an arbitrary element in doing this.

Second, we could do a proportion test for testing whether the proportion of RS is equal to that of SR . However, this requires us to treat SR and RS as coming from two independent samples. That is clearly not correct: they derive from a paired sample (we observe the same subject twice when we consider pairs of decision-making problems).

Third, we could do a sign test. For testing the difference in two variables between paired observations, either a paired t test, or a signed rank test, or a sign test are options. For the first, we need to assume that the difference between the two variables is interval and normally distributed; for the second, we can relax that assumption, but still need to assume that the difference is ordinal; for the third, we need not assume anything about the difference, which seems appropriate here.

Fourth, we could in principle estimate functional forms for the probability weighting and value functions. However, the methodology of common consequence ladders is explicitly designed to avoid estimating value functions (conclusions about probability weighting hold whatever the value function looks like, see Sect. 2). Moreover, an important part of the motivation of the paper is that functional forms are often inadequate and that, therefore, a nonparametric method (such as the one we use) is preferred (see Sect. 1).

We therefore test for probability weighting in paired problems using a sign test. We recode $R = 1$ and $S = 0$, so that $RR = 0$, $SS = 0$, $RS = \textit{positive}$ and $SR = \textit{negative}$ (the labels *positive* and *negative* are conventions for this test) and test whether observed RS and SR are such that the null hypothesis of their being equally likely to occur can be rejected. Since our priors allow deviations in either direction, we report p values from a two-sided test.

A secondary interest is whether we can reject the null hypothesis that patterns across paired problems are due to chance (driven by noise alone). This amounts to testing whether choices in one problem are independent of choices in the problem it is paired with. We test for this by cross-tabulating choices in paired problems and using a Chi-square test for establishing if there is a relationship between the choices.

For testing diminishing sensitivity relative to the end points, we focus first on rung IX of the common consequence ladder. For losses, this allows us to assess whether the probability weighting function is particularly steep for probabilities between 0.05 and 0.15. We will below also consider the same for the interval 0–0.1, but our primary focus is on an interval near to but not including 0. The reason for this focus is the well-known observation that the probability weighting function is ill-defined near the end points, notwithstanding these end points serving as reference points (Kahneman and Tversky 1979). Focusing on rung IX in the domain of gains allows us to assess whether the PWF is particularly steep for probabilities between 0.75 and 0.95.

In Table 5, we show the choice patterns in the domain of losses for rungs paired with rung IX on the common consequence ladder. As mentioned, there are four possible combinations when two rungs are considered: twice safe (*SS*), twice risky (*RR*), first safe then risky (*SR*), and first risky then safe (*RS*). *SS* and *RR* are consistent with EUT, while *SR* and *RS* are not and point to nonlinear probability weighting. Therefore, for example, when choosing between the paired lotteries on rung IX in one decision-making problem and on rung VI in another, 50 subjects chose the safe lottery in each choice problem, 68 the risky lottery both times, 45 chose the safe lottery on rung IX and the risky lottery on rung VI, and 21 chose the

Table 5 Common consequence effects between rung IX and other rungs (losses)

Rungs	Responses (number)				Noise alone? (<i>p</i> value)	SR = RS? (<i>p</i> value)	Inference
	SS	RR	SR	RS			
IX–X	57	59	38	30	0.000***	0.3961	$w(0.15) - w(0.05) \approx w(0.1) - w(0)$
IX–VIII	61	68	34	21	0.000***	0.1048	$w(0.15) - w(0.05) \approx w(0.2) - w(0.1)$
IX–VII	56	69	39	20	0.000***	0.0183**	$w(0.15) - w(0.05) > w(0.3) - w(0.2)$
IX–VI	50	68	45	21	0.000***	0.0043***	$w(0.15) - w(0.05) > w(0.4) - w(0.3)$
IX–V	61	67	34	22	0.000***	0.1409	$w(0.15) - w(0.05) \approx w(0.5) - w(0.4)$
IX–IV	54	66	41	23	0.000***	0.0328**	$w(0.15) - w(0.05) > w(0.6) - w(0.5)$
IX–III	53	64	42	25	0.000***	0.0498**	$w(0.15) - w(0.05) > w(0.7) - w(0.6)$
IX–II	46	63	49	25	0.000***	0.0071***	$w(0.15) - w(0.05) > w(0.75) - w(0.65)$
IX–I	42	64	53	25	0.000***	0.0020***	$w(0.15) - w(0.05) > w(0.8) - w(0.7)$

$N = 184$. Response columns show the number of each of the four possible choice patterns between paired lotteries on indicated rungs of the common consequence ladder in the domain of losses. We use a Chi-square test for testing the hypothesis that choices in paired problems are independent, i.e., are due to “noise alone”; *p* values reported are associated with the Pearson χ^2 test statistic. We use a sign test for testing the hypothesis that $SR = RS$; reported *p* values are based on a two-sided test *Denotes rejection of the null at the 10% level, ** at the 5 percent level, and *** at the 1% level

risky lottery on rung IX and the safe lottery on rung VI. We first check whether we can reject the notion that such choice patterns are due to “noise alone” by using a Chi-square test for testing the hypothesis that choices in paired problems are independent; this hypothesis is always rejected at the 1% level (Tables 5 and 7).

Using a two-sided sign test, we next test whether *SR* is significantly different from *RS* for each of the nine steps involving rung IX. In the example of the step from rung VI to rung IX, the likelihood that *SR* and *RS* are due to error is equal to 0.43%; EUT is rejected at the 1% level. As demonstrated in Sect. 2.4, this points to nonlinear probability weighting. The probabilities for which this may be inferred are spelt out in the final column of Table 5. In total, there are six significant common consequence effects: the PWF for losses is less steep for probabilities [0.2, 0.3], [0.3, 0.4], [0.5, 0.6], [0.6, 0.7], [0.65, 0.75], and [0.7, 0.8] than it is for [0.05, 0.15]. Conversely, no evidence is found that the PWF for losses is less steep for [0.05, 0.15] than for any other interval.

We next investigate whether more can be said about the shape of the PWF when considering all possible common consequence effects for losses (Table 6). Of the 45 possible steps between rungs, 10 point to statistically significant probability weighting. Nine of these point to particular steepness of the PWF for probabilities [0, 0.1] and [0.05, 0.15], and are thus all consistent with concavity of the PWF arising from diminishing sensitivity relative to the end point of $p = 0$. The tenth significant common consequence effect suggests that the PWF is steeper for [0.4, 0.5] than for [0.7, 0.8], but is only significant at the 10% level. Since there are 45 comparisons, this effect is too weak to base a firm conclusion on.

For losses, we thus find strong evidence for a PWF that is relatively steep near $p = 0$. In marked contrast, for gains, we find no evidence that the PWF is relatively

Table 6 All possible common consequence effects (losses)

	PWF steeper for row than column probabilities?								
	0.05–0.15	0.1–0.2	0.2–0.3	0.3–0.4	0.4–0.5	0.5–0.6	0.6–0.7	0.65–0.75	0.7–0.8
0–0.1	ns	ns	ns	s**	ns	ns	ns	s*	s**
0.05–0.15		ns	s**	s***	ns	s**	s**	s***	s***
0.1–0.2			ns	ns	ns	ns	ns	ns	ns
0.2–0.3				ns	ns	ns	ns	ns	ns
0.3–0.4					ns	ns	ns	ns	ns
0.4–0.5						ns	ns	ns	s*
0.5–0.6							ns	ns	ns
0.6–0.7								ns	ns
0.65–0.75									ns

s indicates that the PWF is steeper, *ls* that it is less steep for the row than for the column interval of probabilities; *ns* that the null hypothesis of equal steepness for the two intervals cannot be rejected. Each comparison is tested using a two-sided sign test that $SR = RS$ for the relevant two rungs on the common consequence ladder, as explained in the main text. * denotes rejection of the null at the 10% level, ** at the 5 percent level, and *** at the 1% level

steep near $p = 1$. There are no significant common consequence effects involving rung IX (Table 7), which allows us to investigate the relative steepness of the PWF for [0.75, 0.95]. Likewise, no evidence is found for the PWF being steeper than elsewhere for the interval [0.8, 1.0] (Table 8).

There is some tentative evidence for the PWF for gains being relatively flat in the middle, which is consistent with an inverse S shape. It is less steep for [0.35, 0.55] than it is for [0.6, 0.8] and [0.7, 0.9] and less steep for [0.45, 0.65] than for [0.7, 0.9] (Table 7). However, the dominant impression for gains is the marked absence of probability weighting.

In sum, probability weighting in the domain of losses is pronounced and points to a relatively steep PWF near $p = 0$, which suggests diminishing sensitivity near that probability. This is consistent with an inverse S shape. No strong evidence is found for probability weighting in the domain of gains, but such evidence that is found is also consistent with an inverse S shape.

4.3 Multivariate analysis of nonlinear probability weighting

We next report on correlates of probability weighting. We investigated these by classifying subjects according to their choice patterns across the ten rungs of the common consequence ladder (cf. Bleichrodt and Pinto (2000)), and running probit regressions in which the dependent variable is a dummy that captures one category of the classification. In the domain of gains, 35% of subjects are *EUT consistent* (19% chose always the risky lottery, 16% always safe). We classify the probability weighting of 30% as *dominantly convex*, because the number of times they chose

Table 7 Common consequence effects between rung IX and other rungs (gains)

Rungs	Responses (Number)				Noise alone? (<i>p</i> -value)	SR = RS? (<i>p</i> -value)	Inference
	SS	RR	SR	RS			
IX–X	61	63	30	30	0.000***	1.0000	$w(0.95) - w(0.75) \approx w(1) - w(0.8)$
IX–VIII	67	64	24	29	0.000***	0.5831	$w(0.95) - w(0.75) \approx w(0.9) - w(0.7)$
IX–VII	61	65	30	28	0.000***	0.8957	$w(0.95) - w(0.75) \approx w(0.85) - w(0.65)$
IX–VI	66	61	25	32	0.000***	4270	$w(0.95) - w(0.75) \approx w(0.8) - w(0.6)$
IX–V	58	62	33	31	0.000***	0.9007	$w(0.95) - w(0.75) \approx w(0.75) - w(0.55)$
IX–IV	63	66	28	27	0.000***	1.0000	$w(0.95) - w(0.75) \approx w(0.7) - w(0.5)$
IX–III	63	72	27	21	0.000***	0.4709	$w(0.95) - w(0.75) \approx w(0.65) - w(0.45)$
IX–II	60	64	30	29	0.000***	1.0000	$w(0.95) - w(0.75) \approx w(0.6) - w(0.4)$
IX–I	53	65	38	28	0.000***	0.2678	$w(0.95) - w(0.75) \approx w(0.55) - w(0.35)$

$N = 184$. Responses columns show the number of each of the four possible choice patterns between paired lotteries on indicated rungs of the common consequence ladder in the domain of gains. We use a Chi-square test for testing the hypothesis that choices in paired problems are independent, i.e., are due to “noise alone”; *p* values reported are associated with the Pearson χ^2 test statistic. We use a sign test for testing the hypothesis that $SR = RS$; reported *p* values are based on a two-sided test *Denotes rejection of the null at the 10% level, ** at the 5 percent level, and *** at the 1% level

Table 8 All possible common consequence effects (gains)

	PWF steeper for row than column probabilities?									
	0.4-0.6	0.45-0.65	0.5-0.7	0.55-0.75	0.6-0.8	0.65-0.85	0.7-0.9	0.75-0.95	0.8-1.0	
0.35-0.55	ns	ns	ns	ns	Is**	ns	Is*	ns	ns	ns
0.4-0.6		ns	ns	ns	ns	ns	ns	ns	ns	ns
0.45-0.65			ns	ns	Is*	ns	ns	ns	ns	ns
0.5-0.7				ns	ns	ns	ns	ns	ns	ns
0.55-0.75					ns	ns	ns	ns	ns	ns
0.6-0.8						ns	ns	ns	ns	ns
0.65-0.85							ns	ns	ns	ns
0.7-0.9								ns	ns	ns
0.75-0.95									ns	ns

s indicates that the PWF is steeper, Is that it is less steep for the row than for the column interval of probabilities; ns that the null hypothesis of equal steepness for the two intervals cannot be rejected. Each comparison is tested using a two-sided sign test that $SR = RS$ for the relevant two rungs on the common consequence ladder, as explained in the main text. * denotes rejection of the null at the 10% level, ** at the 5 percent level, and *** at the 1% level

the risky lottery on the first five rungs exceeds that of the last five rungs. It follows that for them *RS* switches dominate between rungs corresponding with lower probabilities and rungs corresponding with higher probabilities, i.e., convexity dominates. For mirroring reasons, 20% are classified as exhibiting probability weighting that is *dominantly concave*, which leaves 15% of subjects unclassified in the domain of gains.

In the domain of losses, 21% of subjects are *EUT consistent* (14% chose always risky; 7% always safe). For classifying the probability weighters, we ignore rungs V and VI, which capture a relatively flat stretch in the middle of the (aggregate) PWF. *Dominant concavity* (36% of subjects) is said to occur when the frequency of *RS* switches between rungs I and IV on one hand and rungs VII–X on the other exceeds that of the frequency of *SR* switches, and *dominant convexity* when it is the other way around (28%). As in the domain of gains, 15% of subjects remain unclassified in the domain of losses.

Tables 9 and 10 report marginal effects for a selection of the probit regressions we ran. In the domain of gains, women are less likely than men to be EUT consistent (Table 9). There is also a small effect of the order in which the lottery choice was presented on the likelihood of EUT consistency. In the domain of losses, risk aversion is positively and significantly correlated with EUT consistency, which (not reported in the table) is driven by its significant correlation with the “always safe” choice pattern.

Neither in the domain of gains nor in the domain of losses were any significant correlates found of the choice patterns dominantly convex and dominantly concave. The reason for this could well be that these classifications allow for a wide variety of choice patterns. We therefore zoomed in on probability weighting near the end points of $p = 0$ in the domain of losses and $p = 1$ in the domain of gains, i.e., comparisons involving rungs IX and X. Recall that, in the domain of gains, we do not find evidence for nonlinear probability weighting near $p = 1$ for the sample as a whole (Table 8). By contrast, we find that convexity near $p = 1$ is more common among traditional farmers (Table 10). In the domain of losses, for the sample as a whole, we found strong evidence for concavity near $p = 0$ (Table 6). By contrast, we find that convexity near $p = 0$ is more common among traditional farmers (Table 10). We will reflect on these contrasting findings for traditional and other farmers in Sect. 5.

5 Discussion and conclusion

Our findings are usefully summarised in terms of our hypotheses (see Sect. 2.2). Hypothesis 1 postulates diminishing sensitivity of the PWF relative to $p^* = 0$, which we investigated and confirm in the domain of losses. Hypothesis 2 postulates diminishing sensitivity of the PWF relative to $p^* = 1$, which we investigated but cannot confirm in the domain of gains. Hypotheses 3 and 4 postulate diminishing sensitivity relative to reference probability p^* , $0 < p^* < 1$. Hypothesis 3 states that the PWF is concave for probabilities $(p^y, 1]$. For gains, this is rejected, as those rejections of expected utility theory we find are all consistent with convexity for

Table 9 EUT consistent: regression analysis

Variable	Dependent variables	
	EUT consistent(gains)	EUT consistent (losses)
Traditional farmer (=1)	-0.07 (0.08)	0.02 (0.07)
Wealth index	0.00 (0.02)	-0.00 (0.01)
Risk aversion (CRRA)	0.02 (0.02)	0.04** (0.02)
Female (=1)	-0.25*** (0.08)	0.04 (0.06)
Years of schooling	-0.00 (0.01)	0.01 (0.01)
Age	0.00 (0.00)	0.00 (0.00)
Harvest failed in past 5 years (=1)	0.03 (0.10)	-0.04 (0.08)
Experienced severe flooding in past 5 years (=1)	0.10 (0.10)	0.01 (0.08)
Experienced severe drought in past 5 years (=1)	-0.08 (0.14)	-0.07 (0.13)
Order 1–10 (=1)	-0.14* (0.08)	-0.04 (0.06)
Experimenter team (=1 for team 1, 0 for team 2)	-0.18 (0.17)	-0.15 (0.14)
Control question correct (=1)	0.05 (0.11)	0.16 (0.06)
<i>N</i>	183	184
Pseudo R^2	0.12	0.12

Probit regression, reporting marginal effects; robust standard errors are in parentheses; session fixed effects were used (not reported in the table). ***, **, and * indicate two-sided significance levels at 1%, 5%, and 10%, respectively

[0.35, 1]. Moreover, convexity of the PWF approaching $p = 1$ is found to be particularly pronounced for traditional farmers. Hypothesis 4 states that the PWF is convex for probabilities $[0, p^y)$, which we reject for the aggregate sample (as implied by the confirmation of hypothesis 1). However, for traditional farmers, convexity of the PWF near $p^* = 0$ is more likely than for others.

How do these findings relate to the previous literature, and what do they suggest? The absence of support in the aggregate sample for a reference probability of $p^* = 1$ in the domain of gains is out of line with the bulk of lab studies in Western countries, as well as with quite a few lab-in-the-field studies in developing

Table 10 Probability weighting near $p = 0$ and $p = 1$: regression analysis

Variable	Dependent variables	
	Convex near $p = 1$ (gains)	Convex near $p = 0$ (losses)
Traditional farmer (=1)	0.13** (0.07)	0.18** (0.09)
Wealth index	0.02* (0.01)	0.00 (0.02)
Risk aversion (CRRA)	0.01 (0.01)	-0.03 (0.03)
Female (=1)	-0.00 (0.05)	-0.06 (0.08)
Years of schooling	0.00 (0.00)	-0.02 (0.01)
Age	0.00 (0.00)	-0.01* (0.00)
Harvest failed in past 5 years (=1)	0.02 (0.06)	0.17* (0.09)
Experienced severe flooding in past 5 years (=1)	-0.04 (0.07)	-0.12 (0.10)
Experienced severe drought in past 5 years (=1)	-0.08 (0.11)	0.15 (0.13)
Order 1–10 (=1)	-0.10* (0.05)	-0.07 (0.08)
Experimenter team (=1 for team 1, 0 for team 2)	-0.06 (0.12)	0.25 (0.19)
Control question correct (=1)	0.04 (0.05)	-0.08 (0.15)
N	176	184
Pseudo R^2	0.16	0.10

Probit regression, reporting marginal effects; robust standard errors are in parentheses; session fixed effects were used (not reported in the table). ***, **, and * indicate two-sided significance levels at 1%, 5%, and 10%, respectively

countries. The dominant probability weighting found in Western labs finds diminishing sensitivity relative to $p^* = 1$; see the reviews of the many studies of probability weighting in Wu and Gonzalez (1996, 1998), Prelec (1998), Gonzalez and Wu (1999), Starmer (2000), Sugden (2004), Stott (2006), Van de Kuilen and Wakker (2011), Fehr-Duda and Epper (2012). The evidence on probability weighting in developing countries is more limited. Imposing Prelec (1998)'s one-parameter functional form, Tanaka et al. (2010) find evidence for an inverse S-shaped PWF in a rural sample from Vietnam. Adopting the methodology developed by Tanaka et al. (2010), Liu (2013) also finds evidence for an inverse

S-shaped PWF, in a sample of Chinese cotton farmers. Eliciting certainty equivalents of binary lotteries, and estimating parameters of Prelec (1998)'s two-parameter functional form, Vieider et al. (2018) find evidence for an inverse S-shaped PWF for a sample of the general population in rural Ethiopia.¹⁷ An inverse S implies a reference probability of $p^* = 1$, so we do not find evidence supporting the lab-in-the field studies for developing countries that find this shape of the PWF. One reason for this could be that imposing a functional form for the entire domain of a PWF may suggest probability weighting for certain ranges of probabilities that would not be found if the investigations zoomed in on that range. For example, estimating parameters of one of Prelec (1998)'s PWFs for a sample in which in reality only diminishing sensitivity near $p^* = 0$ existed and not near $p^* = 1$ would yield estimates consistent with an inverse S and therefore erroneously suggest convexity near 1.0 (which implies diminishing sensitivity near $p^* = 1$).

For that reason, we adopted a methodology in this study that enables more-refined tracking of the PWF: common consequence ladders. It is similar to but much more-refined than Humphrey and Verschoor (2004a, 2004b), who, in five rural samples from Ethiopia, Uganda and India find evidence for an S-shaped PWF. (They investigate gains only, not losses.) As in our study, they use common consequence shifts to investigate nonlinear probability weighting. Unlike in our study, the probabilities they consider are limited to 0, 0.25, 0.5, 0.75, and 1. The contrast between their studies and ours is striking, given the similarity in methods (common consequence ladders) and subjects (poor rural farmers, including in Uganda). They find that the PWF for gains is steep for the middle range of probabilities (in the vicinity of 0.5), whereas we find that it is flat for that range. A hypothesis for explaining the difference is that their particular selection of probabilities made the one in the middle, 0.5, prominent so that it became a reference probability. A reference probability of $p^* = 0.5$ gives rise to an S-shaped PWF, so one that is particularly steep for the middle range of probabilities. This hypothesis may be worth testing in future research.

The reference probability of $p^* = 0$ that we find in the domain of losses is consistent with the literature on probability weighting that included losses, all conducted in Western labs (Tversky and Kahneman 1992; Abdellaoui 2000; Etchart-Vincent 2004; Abdellaoui et al. 2005). The weak evidence we find for an additional reference probability of $p^* = 0.5$ in the domain of losses is at odds with that literature. The evidence is not strong enough to base a firm conclusion on, so may require a higher powered test.

Our findings are plausible in the context of livelihoods strategies of farmers in the rural areas of developing countries. These areas are characterised by frequent shocks that threaten livelihoods, such as droughts, pests and diseases, and so forth (Fafchamps 2003). Our study area is no exception (see Sect. 4.1). These shocks notwithstanding, farmers in our study area do invest: 65% of farmers buy fertiliser for the growing of their crops; 31% grow cash crops, despite great income volatility and sizeable probabilities of investment losses (Verschoor et al. 2016, p.140).

¹⁷ Using a similar approach, l'Haridon and Vieider (2018) find evidence for an inverse S in student samples from 30 countries, including many developing countries.

Persistent probability weighting for a farmer who invests every agricultural season would be surprising, since it causes profits to be lower than they could have been. In other words, probability weighting is a bias that an experienced decision-maker under risk is unlikely to indulge in. Our finding of a marked absence of probability weighting in the domain of gains is consistent with this insight. So is our finding that such probability weighting persists among traditional farmers, who are not in the habit of investing.

The pronounced probability weighting we find in the domain of losses is striking in the light of its marked absence in the domain of gains. We find strong evidence for concavity of the PWF that is consistent with diminishing sensitivity relative to $p^* = 0$. To help interpret this, it is helpful to appreciate the reality of agricultural investment in the study area. The inability to recoup an investment may spell the end of the road for smallholder farmers, since Centenary Rural Development Bank (the main source of agricultural credit in the area) may claim a farmer's land. A nonzero probability that an agricultural investment results in a loss is therefore a very serious matter. Finding diminishing sensitivity relative to $p^* = 0$ in the domain of losses (i.e., caring particularly when the probability of a loss starts to deviate from zero) is therefore less surprising than diminishing sensitivity relative to $p^* = 1$ in the domain of gains would have been for experienced decision-makers under risk. The contrasting finding for traditional farmers is again telling. For them, we find convexity, not concavity, of the PWF for small probabilities. These farmers are not in the habit of investing, so this finding is not likely to reflect their attitudes towards losses that may result from agricultural investment. It may well reflect a more general attitude to uncertainty that stems from farming in a particularly hazardous environment. Every few years a drought will strike, or pests and diseases will destroy the harvest. A probability of zero of no losses is therefore an abstract concept, which may help to explain that small probabilities are treated as not that dissimilar from zero and hence the relative flatness of the PWF in the vicinity of zero.

To conclude, in this study, we investigated the probability weighting habits of farmers from eastern Uganda. Previous studies of risky choice in developing countries have not separately investigated probability weighting for losses, but some studies for Western labs have. We find evidence of a reference probability of 0 in the domain of losses. A reference probability of 0 is consistent with an inverse S-shaped probability weighting function. In that respect, our findings are similar to the bulk of the evidence from Western labs. We do not find evidence for a reference probability of 1 in the domain of gains (apart from for traditional farmers, about one third of our sample). In that respect, our findings are at odds with most evidence from Western labs as well as with earlier studies of probability weighting in developing countries. We interpret this as evidence for experience of risky choice diminishing the tendency to transform probabilities nonlinearly into decision weights, although less so in the domain of losses.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abdellaoui, M. (2000). Parameter-free elicitation of utility and probability weighting functions. *Management Science*, *46*(11), 1497–1512.
- Abdellaoui, M., Driouchi, A., & l'Haridon, O. (2011). Risk aversion elicitation: Reconciling tractability and bias minimization. *Theory and Decision*, *71*, 63–80.
- Abdellaoui, M., Vossman, F., & Weber, M. (2005). Choice-based elicitation and decomposition of decision weights for gains and losses under uncertainty. *Management Science*, *51*(9), 1384–1399.
- Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'École Américaine. *Econometrica*, *21*(4), 503–546.
- Blavatsky, P. R. (2006). Axiomatization of a preference for most probable winner. *Theory and Decision*, *60*, 17–33.
- Bleichrodt, H., & Pinto, J. L. (2000). A parameter-free elicitation of the probability weighting function in medical decision analysis. *Management Science*, *46*, 1485–1496.
- Dave, C., Eckel, C. C., Johnson, C. A., & Rojas, C. (2010). Eliciting risk preferences: When is simple better? *Journal of Risk and Uncertainty*, *41*(3), 219–243.
- Etchart-Vincent, N. (2004). Is probability weighting sensitive to the magnitude of consequences? An experimental investigation on losses. *Journal of Risk and Uncertainty*, *28*(3), 217–235.
- Etchart-Vincent, N., & l'Haridon, O. (2011). Monetary incentives in the loss domain and behavior toward risk: An experimental comparison of three reward schemes including real losses. *Journal of Risk and Uncertainty*, *42*, 61–83.
- Fafchamps, M. (2003). *Rural poverty, risk and development*. Cheltenham: Edward Elgar.
- Fehr-Duda, H., & Epper, T. (2012). Probability and risk: Foundations and economic implications of probability-dependent risk preferences. *Annual Review of Economics*, *4*, 567–593.
- Filmer, D., & Pritchett, L. H. (2001). Estimating wealth effects without expenditure data—or tears: An application to educational enrolments in states of India. *Demography*, *38*(1), 115–132.
- Gneezy, U., & Potters, J. (1997). An experiment on risk taking and evaluation periods. *The Quarterly Journal of Economics*, *112*(2), 631–645.
- Gonzalez, R., & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology*, *38*, 129–166.
- Harrison, G. W., Humphrey, S. J., & Verschoor, A. (2010). Choice under uncertainty: Evidence from Ethiopia, India and Uganda. *The Economic Journal*, *120*(543), 80–104.
- Hey, J. D., & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica*, *62*, 1291–1326.
- Humphrey, S. J., & Verschoor, A. (2004a). Decision-making under risk among small farmers in East Uganda. *Journal of African Economies*, *13*(1), 44–101.
- Humphrey, S. J., & Verschoor, A. (2004b). The probability weighting function: Experimental evidence from Uganda, India and Ethiopia. *Economics Letters*, *84*, 419–425.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–291.
- Koszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, *121*(4), 1133–1165.
- Koszegi, B., & Rabin, M. (2007). Reference-dependent risk attitudes. *The American Economic Review*, *97*(4), 1047–1073.
- l'Haridon, O., Vieider, F. M. (2018). All over the map: A worldwide comparison of risk preferences. *Quantitative Economics*, forthcoming.

- Liu, E. M. (2013). Time to change what to sow: Risk preferences and technology adoption decisions of cotton farmers in China. *The Review of Economics and Statistics*, 95(4), 1386–1403.
- Neilson, W., & Stowe, J. (2002). A further examination of cumulative prospect theory parameterizations. *The Journal of Risk and Uncertainty*, 24(1), 31–46.
- Prelec, D. (1998). The probability weighting function. *Econometrica*, 66(3), 497–527.
- Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization*, 3, 323–343.
- Segal, U. (1989). Anticipated utility: A measure representation approach. *Annals of Operations Research*, 19, 359–373.
- Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature*, 38(2), 332–382.
- Stott, H. P. (2006). Cumulative prospect theory's functional menagerie. *Journal of Risk and Uncertainty*, 32(2), 101–130.
- Sugden, R. (2004). Alternatives to expected utility: Foundations. In S. Barberà, P. J. Hammond, & C. Seidl (Eds.), *Handbook of utility theory* (Vol. 2, pp. 685–755). Extensions Boston: Springer.
- Takahashi, T. (2011). Psychophysics of the probability weighting function. *Physica A*, 390, 902–905.
- Tanaka, T., Camerer, C. F., & Nguyen, Q. (2010). Risk and time preferences: Linking experimental and household survey data from vietnam. *The American Economic Review*, 100(1), 557–571.
- Thaler, R. H., & Johnson, E. J. (1990). Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management Science*, 36(6), 643–660.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5, 297–323.
- Van de Kuilen, G., & Wakker, P. P. (2011). The midweight method to measure attitudes toward risk and ambiguity. *Management Science*, 57(3), 582–598.
- Verschoor, A., D'Exelle, B., & Perez-Viana, B. (2016). Lab and life: Does risky choice behaviour observed in experiments reflect that in the real world? *Journal of Economic Behavior and Organization*, 128, 134–148.
- Vieider, F. M., Beyene, A., Bluffstone, R., Dissanayake, S., Gebreegziabher, Z., Martinsson, P., et al. (2018). Measuring risk preferences in rural Ethiopia. *Economic Development and Cultural Change*, 66(3), 417–446.
- Vieider, F. M., Martinsson, P., Nam, P. K., & Truong, N. (2019). Risk preferences and development revisited. *Theory and Decision*, 86, 1–21.
- Wakker, P. P. (2010). *Prospect theory for risk and ambiguity*. Cambridge: Cambridge University Press.
- Wu, G., & Gonzalez, R. (1996). Curvature of the probability weighting function. *Management Science*, 42(12), 1676–1690.
- Wu, G., & Gonzalez, R. (1998). Common consequence conditions in decision making under risk. *Journal of Risk and Uncertainty*, 16, 115–139.
- Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica*, 55, 95–115.
- Yesuf, M., & Bluffstone, R. A. (2009). Poverty, risk aversion, and path dependence in low-income countries: experimental evidence from Ethiopia. *American Journal of Agricultural Economics*, 91(4), 1022–1037.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.