The Devil is in the Details:

Risk Preferences, Choice List Design, and Measurement Error

Stein T. Holden and Mesfin Tilahun

Norwegian University of Life Sciences
Centre for Land Tenure Studies

The Devil is in the Details:

Risk Preferences, Choice List Design, and Measurement Error[1]

By

Stein T. Holden[1] and Mesfin Tilahun[1,2]
[1]School of Economics and Business
Norwegian University of Life Sciences
P. O. Box 5003, 1432 Ås, Norway.
[2]Mekelle University, Mekelle, Tigray, Ethiopia.
Email: stein.holden@nmbu.no; mesfin.tilahun.gelaye@gmail.com

Abstract

We use a field experiment to estimate the risk preferences of 945 youth and young adult members of 116 rural business groups organized as primary cooperatives in a semi-arid risky environment in northern Ethiopia. Multiple Choice Lists with binary choices between risky prospects and varying safe amounts are used to identify the certainty equivalent for each risky prospect. Rank Dependent Utility Models with alternatively Wilcox' (2011) Contextual Utility or Busemeyer and Townsend (1992, 1993) Decision Field Theory heteroskedastic error specifications are used to estimate risk preference parameters and parametrized model noise. The study aims to a) assess potential biases associated with Choice List design; b) assess a time-saving elicitation method; c) inspect the explanatory power of the predicted risk preference parameters for respondents' investment, income and endowment variables; d) assess how the predictive power is associated with model noise and the addition of two low probability high outcome risky prospects that may help to capture utility curvature more accurately. Substantial risk parameter sensitivity to Choice List design was detected. The rapid elicitation method appears attractive as it facilitates use of a larger number of Choice Lists with variable attributes although it is sensitive to bias due to random error associated with randomized starting points. The addition of the two Choice Lists with low probability high

outcomes substantially enhanced the explanatory power of the predicted risk preference parameters and resulted in substantially higher estimates of the utility curvature parameter (higher risk aversion).

Key words: Risk preferences, rank dependent utility, probability weighting, measurement error, predictive power, field experiment, Ethiopia.

JEL codes: C90; C93; D14, D81; D90.

1. Introduction

Risk and risk preferences are fundamentally important for the survival strategies of resource-poor people living in risky environments. Risk preferences are latent variables that may be revealed through experiments or survey questions. However, it is challenging to identify and measure such preferences accurately (Choi et al. 2014; Loomes & Pogrebna 2014); Dohmen et al. 2018). No consensus exists on the best way to measure risk preferences and there is large variation in the methods used (Zhou and Hey 2018). Based on expected utility theory (EUT) risk preferences are associated with the curvature of the utility function (Arrow 1965; Pratt 1964). Under rank dependent utility (RDU) (Quiggin 1982) and cumulative prospect theory (CPT) (Tversky and Kahneman 1992) risk preferences are also associated with probability weighting, while CPT also includes loss aversion as a specific form of risk preference associated with losses.

Potential measurement errors associated with measuring the alternative representations of risk preferences have recently received attention by researchers and may be an important reason for the large heterogeneity in findings (Chuang and Schechter 2015; Schildberg-Hörisch 2018; Dohmen et al. 2018). Especially cognitive ability and the design of the different tools used for eliciting risk preferences have been subject to recent research and have revealed new insights (Andersson et al. 2016; Benjamin et al. 2013; Burks et al. 2009; Dohmen et al. 2010; 2018; Oechler et al. 2009; Vieider 2018). Earlier studies have found that there can be order effects and these may be due to learning and fatigue when a large number of cognitively demanding questions have to be answered (Harrison et al. 2005; Andersen et al. 2006). The responses to Choice Lists (CLs) may also depend on the magnitude levels and the distribution of the part of the CLs that is varying (Andersen et al. 2006). There can be anchoring effects and bias towards the middle for psychological reasons in such lists if the respondents have to go through the whole lists from the top to the bottom (or bottom-up) (Andersen et al. 2006). Recently, more focus has been on the possibility of bias due to random choice in CLs (Andersson et al. 2016). Such random choice may cause biases if the correct choice is close to one end of the list because of the implied censoring of the error distribution. Andersson et al. (2016) found that cognitive ability was confounded with such errors and resulted in biased estimates of risk preferences but the extent and direction of such bias depended on the design of the CL and the position of the risk-neutral task in the list.

We aim to contribute to this literature based on a field experiment where we demonstrate that the estimated risk preference parameters and their predictive power can be highly sensitive to the design of the CL included in the field experiment. The field experiment used a sample of 945 youth and young adults that have joined 116 youth business groups to develop a joint business as a complementary livelihood strategy in a risky environment in rural Ethiopia. The respondents have limited education with a median of six years completed education.

We assess how the order of the CLs (we randomized the order), the position of the (randomly) chosen starting point in each CL, and the position of the risk-neutral row in the CLs affected the estimated noise parameter as well as the estimated risk preference parameters. This allowed us to critically examine possible design errors and their bias implications and these lessons can be helpful for improving future designs. This is our first contribution.

We test a rapid elicitation method that may hold the potential to reduce some types of bias in addition to being time saving, and inspect its sensitivity to random error bias. Earlier studies have revealed problems with anchoring, bias towards the middle and fatigue associated with exposing respondents to MCLs involving a large number of questions (Andersen et al. 2006). We use series of paired gambles where each requires a choice between a risky prospect and varying sure amounts to identify the certainty equivalent for the risky prospect. Multiple Choice Lists (MCLs) with varying probabilities for high and low outcomes for the risky prospects are the organizing tool for the experimental enumerators but are not presented as lists to the respondents. Our approach with randomly selecting a starting row in each choice list (CL) and quickly narrowing in towards a switch point aimed to reduce these sources of bias and test potential biases associated with our approach. This is the second contribution of this paper.

Our third contribution relates to CL and MCL design. We build on the approach of l'Haridon and Vieider (2019) and Vieider et al. (2018) and systematically vary the probabilities across the CLs and only vary the sure amounts within lists. We aimed thereby to get good estimates of the probability weighting function, and particularly so in the area with low and medium probabilities for the low outcome in the risky prospect. The semi-arid environment with frequent droughts and other shocks in which the respondents live and try to establish their livelihoods is likely to have formed their risk preferences in this probability region. At the same time, their joint businesses require them to make risky investments such that their risk preferences are expected to affect both

individual and group livelihood strategies. Our specific contribution is that we added two choice lists in the region with low probability high outcomes in the risky prospect to better map the probability weighting function in this probability region and to get more variation in the outcome levels that were within our budget and thereby get better estimates of the utility curvature. The estimations without and with these two additional CLs and with controls for CL design characteristics, random starting point and order of CLs and their impact on the estimated risk preference parameters is the third contribution. For this we use two alternative approaches to control for the context of CLs, Wilcox' (2011; 2015) Contextual Utility (CU) and Busemeyer and Townsend's (1992; 1993) Decision Field Theory (DFT) that use the distance between the upper and lower limits of the risky prospects as reference points. The advantage of these is that they relax the homoscedasticity assumption that is associated with the standard binary choice models (Wilcox 2011; 2015).

Our fourth contribution is that we assess the explanatory power of the alternative predicted measures of risk preferences in form of their degree of correlation with a number of respondent investment, income and endowment variables. We utilize the random CL design characteristics as instruments to predict the risk preference parameters and parametrized noise. We use the degree of correlation between the predicted risk preference parameters and the real economic variables as indicators of the predictive power in line with economic theory without imposing any strong causality assumptions. Inconsistencies in signs and low and insignificant correlations are potentially indicators of weak designs and estimation methods. Our main finding is that the inclusion of the two extra CLs with low probability high outcome risky prospects substantially improved the predictive power of the estimated risk preferences.

We regard the study as exploratory but think that a number of valuable insights were gained that can help ourselves and others to design better experimental instruments for future field experiments in challenging environments with low resource respondents.

2. Literature review

The optimal choice of experimental approach to elicit respondents' risk preferences depends on the respondents' characteristics, such as cognitive ability, their motivation and willingness to set aside time for the experiments, as well as researchers' resources and objectives. Much of this experimental research was implemented with university students as respondents in laboratory

experiments. It is not obvious that the best tools used in such contexts also are the best tools for eliciting risk preferences from a broader sample of respondents that has less education. We first review literature to gain insights about how context and cognitive ability may constrain or influence measures of risk preferences. Second, we review how experimental designs may influence the measurement of risk preferences.

2.1. Cognitive Limitations, Measurement Error and Potential Biases

Burks et al. (2009) and Benjamin et al. (2013) found that individuals with lower cognitive ability respond less consistently in risk experiments. Dave et al. (2010) assessed simpler versus more complex and sophisticated devices for elicitation of risk preferences. They found that the simpler devices were preferable for respondents with lower numeracy skills. A more sophisticated tool resulted in more noise for respondents with lower numeracy skills but gave superior predictive power for respondents with better numeracy skills. Chuang and Schechter (2015) found weak correlations between choices in risk experiments repeated over time for the same respondents in field experiments in rural Paraguay. They proposed that this may be due to the low level of education of their sample. Choi et al. (2014) proposed that lower decision-making ability may be associated with lower decision-making quality. It is, however, challenging to separate the true preferences from decision-making ability and quality (Dohmen et al. 2018; Zhou and Hey 2018). Choi et al. (2014) proposed to measure the decision-making quality by the degree of consistency of choices with economic rationality in terms of completeness and transitivity. They assessed the correlations between responses in lab experiments in form of heterogeneity in decision-making ability and quality, and how the variations in these are systematically correlated with a socioeconomic characteristic in form of wealth. They found an economically large and statistically significant correlation between consistency in the experiment and household wealth. A standard deviation difference in consistency score for the person in the household in charge of financial matters was associated with a 15-19 percent more household wealth.

Risk preference experiments take many formats and there is yet no consensus on what is the best design of risk preference elicitation tasks (Zhou and Hey 2018). Multiple Choice Lists[2] (MCLs) represent a popular format and include a number of variants from the famous Holt and Laury (2002) format with choice between a less risky and more risky lottery where the probabilities of

---

[2] Often called Multiple Price Lists but we think Choice List is a better name for it than a Price List.

good and bad outcomes change systematically through the list. A second MCL approach for risk preference elicitation compares a risky prospect with a constant probability with a safe option and varies the safe option to identify the certainty equivalent amount (Abdellaoui 2011; Vieider et al. 2018). A third MCL option varies the high or low outcome in the risky prospect but keeps the probability constant in each choice list (e.g. Tanaka et al. 2010).

An increasing number of studies have compared alternative designs and found that the elicited risk preferences are sensitive to the design characteristics. For example, Bruner (2009) found that the elicited risk preferences are sensitive to whether one varies the probabilities or the outcomes in the paired choices in the MCLs. Hey et al. (2009) compared four methods of eliciting risk preferences including pairwise choices, certainty equivalent elicitation through willingness to accept second price auctions, certainty equivalent elicitation through willingness to pay second price auction, and the certainty equivalent elicitation using the Becker-DeGroot-Marschak mechanism. They found that the pairwise choice approach was associated with less noise and less bias than the other methods. Csermely and Rabas (2016) made a comprehensive assessment of MCL risk preference eliciting approaches with a within-subject design using nine different variants of MCLs while using the Holt and Laury (2002) design as the benchmark. They found that the paired game changing the high reward performed the best. Csermely and Rabas (2016) found that varying the potential minimum payoff resulted in more risk averse behavior while including certainty equivalents stimulated risk taking. Similarly, Freeman et al. (2019) found that MCLs with a risky and a sure option lead to significantly more risky choices.

Lévy-Garboua et al. (2012) tested the effect of sequential versus simultaneous presentations of the Holt and Laury MCL as well as increasing, decreasing and random probabilities in the MCL. They found that the responses were sensitive to these variations. Inconsistency errors were more common in sequential than in simultaneous presentations and in decreasing and random than in increasing probabilities of winning. Risk aversion was significantly higher in sequential than in simultaneous framing, in decreasing and random than in increasing probabilities of winning, and when stakes were higher. Similarly, Friedman et al. (2019) found that a monotone sequencing of probabilities and prices across MCLs rather than random order resulted in less risk averse responses.

Bosch-Domènech et al. (2013) assessed the effect of removing some of the rows in the Holt and Laury (2012) MCL. The concluded that removal of good pairs where most subjects choose the risky option resulted in significantly fewer safe choices implying lower risk aversion. They also assessed this for the alternative approach of Abdellaoui et al. (2011) where a risky prospect is compared to sure amounts and found that this MCL was less sensitive to the removal of pairs (no significant effect). It is likely that comparing two risky prospects is cognitively more demanding than comparing one risky prospect with constant probability with varying safe amounts. The latter approach (which we also use in this paper) may therefore be less sensitive to such framing effects.

Drichoutis and Lusk (2012; 2016) indicated that the Holt and Laury (2002) MCL is less suitable to identify utility curvature than probability weighting and is confounding these. They suggested it may do a better job in identifying the probability weighting function because it varies the probabilities within CLs. They proposed and tested an alternative CL, which keeps the probability constant (at 0.5) and varies the payoffs of the risky prospects. It is better at capturing utility curvature when multiple lists cannot be used[3]. Still a single list cannot be used to separate the utility curvature and probability weighting components for each respondent. Tanaka et al. (2010) used three CLs to elicit three parameters, one for utility curvature, one for probability weighting, and one for loss aversion. Errors in elicitation with such an approach also leads to correlated errors due to their exact identification requirement. More CLs than parameters to be estimated are needed to get statistical measures of measurement error. Variants of such MCLs have become common and have also resulted in more focus on measurement errors related to risk preference elicitation.

Another issue is how the MCL information is presented to the respondents. Habib et al. (2017) used the Holt and Laury (2002) game as a base and compared alternative approaches to presenting the MCL information in form of pairwise text versions and alternative graphical pie charts. They hypothesized that the graphical pie charts may convey clearer information and that the respondents would appear less risk averse due to a better ability of calculating expected returns with the graphical presentation (reduced ambiguity for ambiguity averse individuals). Their results supported this hypothesis but the statistical significance was weak due to noise and small samples.

---

[3] Holt and Laury (2002) used multiple lists with varying payoffs.

2.2. Choice List Design and Measurement Error

Poulton and Poulton (1989) and Andersen et al. (2006) found that there is a tendency towards anchoring towards the middle of CLs. With the risk neutral task placed in the middle of a list this could thus also pull towards more risk-neutral responses. With the risk neutral task placed towards one end of the list and a more risk averse task placed in the middle it could lead to a bias in direction of risk aversion also for risk neutral respondents.

Andersson et al. (2016) assessed how limited cognitive skills could contribute to random choice in MCLs and how this in interaction with the CL design could lead to biased estimates of risk preferences. They found that bias due to random choice was more important than bias towards the middle (not significant) in their study. They suggested that the correlations found between cognitive skills and risk aversion found in some earlier studies could be due to measurement errors creating spurious correlations.

Early experiments found that decision makers tend to be more risk averse when a riskless option is available compared to when a certain alternative is unavailable and this has been called the "certainty effect" or "preference for certainty" associated with the common ratio effect (Allais 1953; Camerer 1995).

Callen et al. (2014) used two CLs in a study in Afghanistan and found a "preference for certainty" as a risky prospect was compared to a certain amount in one CL and with another risky prospect in another CL while changing the probabilities for good and bad outcomes of the first risky prospect in both CLs. This result is puzzling because several other studies that use CLs comparing a risky prospect with sure amounts tend to find less risk averse responses than for other CL designs (Csermely and Rabas 2016; Vieider et al. 2018; L'Haridon and Vieider 2019; Freeman et al. 2019). Vieider (2018) proposed that this puzzling reversal is due to the specific design of the two CLs of Callen et al. (2014) and suggested that the "preference for certainty" is driven by random errors by the respondents. He also proposed an alternative design which should result in the opposite "preference for uncertainty" if random error and not "preference for certainty" is the underlying explanation. He implemented both experiments and found what he predicted. He also found indications of reference dependence and loss aversion being invoked in the CLs of Callen et al. (2014) where probabilities were changed instead of sure amounts and this may have resulted in bias towards higher levels of risk aversion being measured.

Freeman et al. (2019) found that embedding a pairwise choice in a CL increases the fraction of subjects choosing the riskier lottery from 23% to 45% when the safer alternative is certain, while this did not affect the choices when the safer alternative is risky. They suggested that the introduction of the CL eliminates the "certainty effect" by introducing the risk from the random incentive mechanism associated with only one task being randomly chosen for real payout.

3. Choice lists and elicitation procedure

3.1. Risk preference experiment

We used a series of Multiple Choice Lists (MCLs) where each choice list (CL) was designed with a risky prospect that did not vary within the CL. We used the certainty equivalent approach. The risky prospect was compared to varying certain amounts to identify the certain amount where the respondents preferred to switch between the risky prospect and the certain amount. The experiment was incentivized by drawing one of the CLs randomly to be played for real and randomly drawing a row in that CL. If the certain amount was chosen for this row, respondents received this amount. If the risky prospect was chosen, the die is used to identify the outcome, see Appendix 2 for details.

CL design

We gave emphasis to careful mapping of the probability weighting function in the area with 5-50% probabilities of bad outcome (low probability negative outcomes) because of our focus on livelihood risks that our sample respondents are exposed to, such as climate risks. The risks of drought for example are within this probability range in this semi-arid area. Other production risks or health risks also fall in this category of low probability bad outcome risks that typically have been associated with risk aversion. The two last series are included to capture the responses to low probability high gain opportunities and to map out that part of the weighting function (although a lower precision can be expected there) and to get a better estimate of the utility curvature by including larger good outcome values.

Table 1 gives an overview of the 12 CLs used in the risk experiment. The order of the CLs was randomized. By including the order of each CL in the estimation, we could test for order effects such as learning and fatigue. In the estimation, we tested for the order effect on the noise parameter as well as the risk preference parameters.

Table 1. Overview of risky prospects in risk game

Choice List	Prob(bad outcome)	Bad outcome, ETB	Good outcome, ETB
1	1/20	0	100
2	1/10	0	100
3	2/10	0	100
4	3/10	0	100
5	5/10	0	100
6	1/20	20	100
7	1/10	20	100
8	2/10	20	100
9	3/10	20	100
10	5/10	20	100
11	15/20	20	300
12	19/20	20	1500

Note: ETB=Ethiopian Birr.

Use of such MCLs may create bias for various reasons as they imply a very large number of binary choices with small differences from row to row in each list. This can make respondents bored and reduce their effort in identifying switch points that accurately reflect their preferences. This could lead to random choices, starting point bias or bias towards the middle (Andersen et al. 2006; Andersson et al. 2016). Freeman et al. (2019) have also found that such CLs with a choice between a risky and a sure option can lead to significantly more risky choices compared to isolated binary choices.

The position of the risk-neutral task row based on linear probability weighting (EUT) varied across the CLs. We tested for potential bias due to the positioning of the risk-neutral row in each CL. This was done for the noise parameter as well as the risk preference parameters.

Elicitation procedure

We used a procedure to reduce the number of questions and reduce the risk of starting point bias and bias towards the middle. This procedure may also reduce the tendency towards higher probability of choice of the risky lottery (Freeman et al. 2019). Our procedure was as follows:

The MCLs were not shown to the respondents but were the guiding tool for the experimental enumerators. The amount of money for the risky prospect (good and bad outcome) for the CL was put on the table in front of the respondents. A 20-sided die was used to explain the probabilities

that varied across CLs. For the initial question in each CL a randomized (in advance by the experimental enumerator) row on the CL was identified as the first binary choice between the row-specific certain amount to be compared with the risky prospect and the given probability for the good versus the bad outcome for the risky prospect in the CL. The respondent answered whether s/he prefers the risky prospect or the certain amount. If the certain amount is preferred, the instructions to the enumerator were to go to the bottom of the list and ask the preference for the risky prospect versus the lowest certain amount. This is likely to lead to a preference for the risky prospect. Then the enumerator was guided to go to the middle between the randomly chosen first row/certain amount and the lowest amount, and so on to rapidly narrow in towards the switch point. This implied that maximum one switch point was identified in each CL. It also implied a more specific focus on a sequence of more isolated binary choices. Random errors in each binary choice may, however, also lead to substantial errors. We tested for starting point bias associated with the random starting row in each CL as a potential source of bias in the first decision in each CL. This was done for the noise parameter as well as the risk preference parameters.

3.2. Sample and data

Table 2 provides distributional information about some key respondent characteristics potentially indicative of cognitive ability and economic indicators that may be correlated with risk preferences. The gender dummy variable was included as quite a few studies have identified a gender difference in risk preferences.

Figures 1a and 1b give an overview of the distribution of the switch points in the first 1-10 CLs and CL 11 and 12. We observe a weakness in the design of CLs 11 and 12 as there is a substantial share (20-25%) having no switch point. This may have resulted in bias in our estimation due to censoring of the error distribution. We attempt to test and partially control for this by specifying the position of the risk neutral row in the CLs. We assess this also by including the probability of bad outcome in the parametrized error specification. The error should increase with this probability. This error may also have reduced the value added of these two CLs in terms or their added contribution to the predictive power. This is one of the issues we assess.

Table 2. Basic sample characteristics.

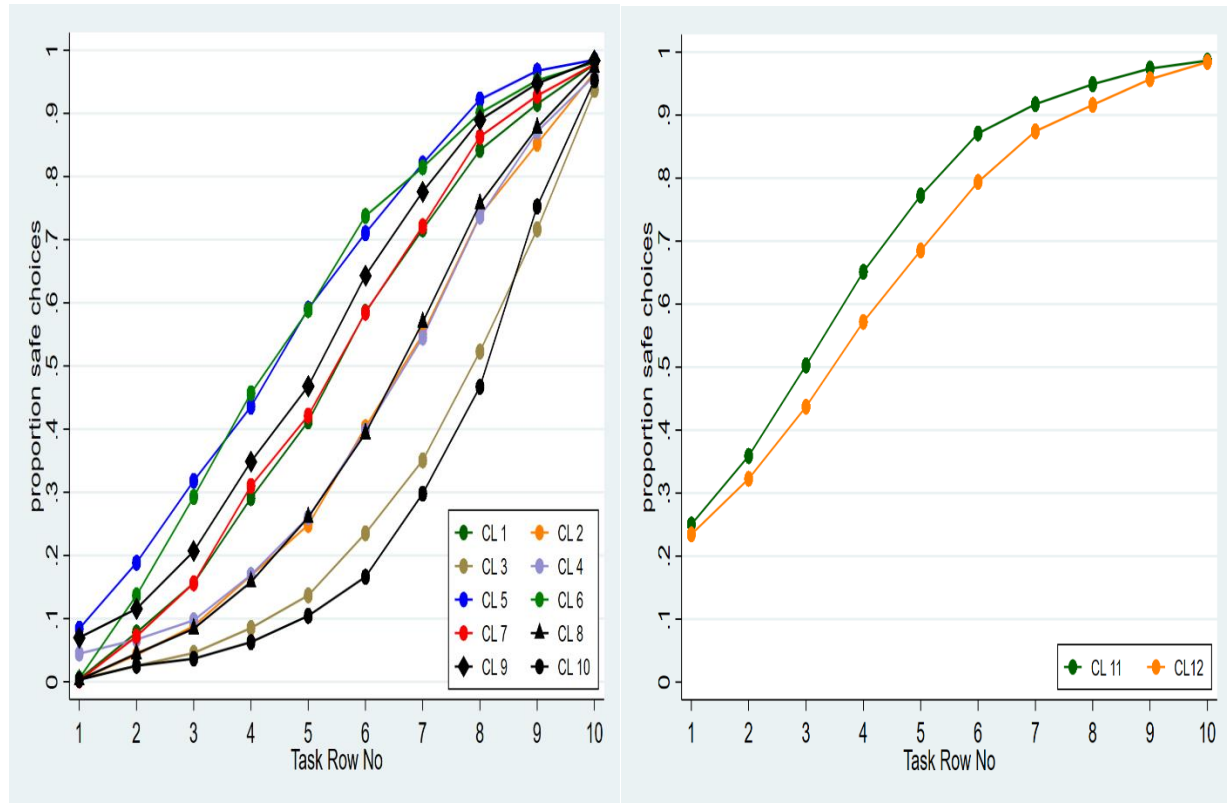| Stats	Male head, dummy	Age, years	Education years	Other individual income, 1000 ETB	Durable assets, number	Investments last year, 1000 ETB	Tropical livestock units
p10	0	19	0	0	0	0.43	0
p25	0	22	2	1.2	0	1.4	0
p50	1	27	6	5	1	4.12	0.7
p75	1	35	9	10	2	10.5	2.1
p90	1	42	10	18	3	20.1	3.3
mean	0.68	29.37	5.38	7.68	1.33	9.42	1.24
se(mean)	0.02	0.32	0.13	0.33	0.05	0.57	0.05
N	945	945	945	945	945	945	945

Figure 1a and 1b. Cumulate switch point distributions for the 1-10 CLs and CL 11 & 12.

4. Estimation methods

4.1. Estimation of risk preference parameters

Each choice of the respondent is between a risky and a safe option. The risky option gives a high outcome (x) with probabilility p and a low outcome (y) with probability 1-p. We call the safe amount s. We place the choice between the risky and safe prospect into a Rank Dependent Utility (RDU) framework (Quiggin 1982). The net utility or value return for a specific risky and a safe option can then be formulated as:

(6) $$\Delta RDU = w(p)u(x) + \left[1 - w(p)\right]u(y) - u(s)$$

where $w(p)$ is the probability weighting function. This model also nests the EU model with linear probability weighting. Since we only allow non-negative outcomes this model is also consistent with Cumulative Prospect Theory (CPT) (Tversky and Kahneman 1992)[4].

We use the two-parameter Prelec 2 probability weighting function (Prelec 1998) to capture the behavioral phenomenon that decision makers typically overweight low probabilities and underweight high probabilities:

(7) $$w(p) = e^{-\beta(-\ln p)^{\alpha}}, \alpha > 0, \beta > 0$$

This is a strictly increasing and continuous function $w(p) : [0,1] \to [0,1]$ with an inverse, w^{-1}, that is also a probability weighting function.

We use a constant partial relative risk aversion (CPRRA) utility function:

(8) $$u(x) = (1-r)^{-1}\left((b+x)^{1-r} - 1\right)$$

where r is the constant partial relative risk aversion (CPRRA) coefficient[5] and b is a base consumption level set at a daily wage rate rather than wealth. This implies that we assume limited

---

[4] L'Haridon and Vieider (2019) elicited probability weighting functions also in the loss domain but found that the shape of the probability weighting function in the loss domain was quite close to that of the gain domain function in a large multi-country sample.

[5] This implies that we assume partial asset integration and relate the risky prospects with a small base consumption level, based on empirical evidence of limited asset integration (Binswanger 1981; Rabin 2000; Wik et al. 2004; Andersen et al. 2008).

asset integration and this is in line with empirical evidence (Binswanger 1981; Rabin 2000) and assumptions made by others (e.g. Andersen et al. 2008).

As our respondents have limited education, they may have problems understanding the games or making calculations correctly in the games, we expect errors in their responses, and such errors may imply violations of consistency. Experimental enumerators may also be a source of error and there may be starting point bias or bias associated with the random order of the CLs. The data from these experiments are therefore noisy and such noise needs to be taken into account in the estimation. Each decision between a risky and a sure amount may thus be subject to such errors. As a first step to factor in the CL design we draw on the contextual utility models of Wilcox (2011; 2015) and the decision field theory (DFT) of Busemeyer and Townsend (1992; 1993).

The suitability of random utility models (RUM) for estimating risk preferences has recently been questioned (Apesteguia and Ballester 2018). Apesteguia and Ballester (2018) claim that non-monotonicity may cause bias and therefore recommend the use of random parameter models (RPM) instead. However, Conte and Hey (2018) question the monotonicity argument and find that RUM performs as well as RPM in a set of simulations with RUM and RPM models. RPM may work well and be preferable in the case of the simpler EU model with only one parameter capturing risk preferences. Vieider (2019, pers. com.) notes that Apesteguia and Ballester's (2018) point does not apply to Certainty Equivalents as long as utility is compared to monetary amounts.

The contextual utility (CU) model of Wilcox (2011) does not suffer from the monotonicity problem in our case with simple prospects and allows joint estimation of utility curvature and a two-parameter probability weighting function and a noise parameter that can be made dependent on prospect and respondent characteristics. The CU model in the case of RDU implies that the probability of the choice of the risky prospect over the safe prospect is:

(9) $$\Pr\left(Risky\right) = F\left(\frac{\lambda \Delta RDU}{u\left(x\right) - u\left(y\right)}\right)$$

This implies a standardization to the utility difference of the high versus low outcomes of the risky prospect, where the safe outcome falls between these, and with weighting through the utility and probability weighting functions. A limitation of this approach may be that it only weighs in the

utility and not the probability weighting. The following approach may therefore be more appropriate in the case of RDU[6]:

The decision field theory (DFT) adjustment based on Busemeyer and Townsend (1992; 1993) and Wilcox (2015) brings the probability weighing into the denominator as well as follows:

$$(10) \qquad \Pr\left(Risky\right) = F\left(\frac{\lambda \Delta RDU}{\left[u(x) - u(y)\right]\sqrt{w(p)\left[1 - w(p)\right]}}\right)$$

Both these approaches have the common property that a higher level of risk aversion will lead to a lower probability of choosing the risky prospect. The DFT approach also has the advantage that as p approaches 0 or 1 the probability of choosing the stochastically dominating alternative approaches certainty (Wilcox 2015).

We estimate these models by specifying them as probit normal probability density functions for respondents i and CLs m and include a heteroskedastic Fechner error ε_{im} with contextual utility:

$$(11) \qquad Probit\left(Risky\right) = \phi\left(\frac{\lambda \Delta RDU_{im}}{\varepsilon_{im}\left[u_{im}(x_{im}) - u_{im}(y_{im})\right]}\right)$$

and the DFT approach with heteroskedastic Fechner error ξ_{im}:

$$(12) \qquad Probit\left(Risky\right) = \phi\left(\frac{\lambda \Delta RDU_{im}}{\xi_{im}\left[u_{im}(x_{im}) - u_{im}(y_{im})\right]\sqrt{w_{im}(p_{im})\left[1 - w_{im}(p_{im})\right]}}\right)$$

The errors allow for within respondent errors in identification of switching points and thereby the CL-level estimates of the weighting function and utility curve parameters. The models are estimated by maximum likelihood for the log likelihood functions for these density functions that are related to the switch point in each CL:

(13)

---

[6] We also tested models with "Contextual Income" where the utility difference was replaced by income difference in the risky prospect. However, these models did not perform as well due to convergence problems when making the noise and risk preference parameters linearly dependent on CL design and respondent characteristics.

$$\ln L\left(z_i, m, \alpha_i, \beta_i, r_i, \theta_k, \xi_{im}\right) = \sum_{im}\left(\left(\ln \Phi\left(\Delta RDU\right) | Choice_{im} = 1\right)\right) + \left(\ln \Phi\left(1 - \Delta RDU\right) | Choice_{ij} = 0\right)$$

We allow the total likelihood to depend linearly on CL design characteristics (m), risk preference parameters (α_i, β_i, r_i), individual observable characteristics (z_i), enumerator dummies (θ_k), and noise that varies systematically across respondents and CLs (ξ_{im}). We cluster standard errors at the individual respondent level.

4.2. Assessment of predictive power of alternative estimated risk preference parameters

We assess the predictive power of the alternative estimates of the risk preference parameters by assessing their correlations with four investment (individual investment during last year), income (individual income during last year) and endowment variables (number of durable assets and livestock endowment measured in Tropical Livestock Units). We represent all of these in the vector Y_i below. Equations 14a-d show the four specifications with the four alternative predicted risk preference parameters as well as the predicted noise parameters. The randomized CL design characteristics have served as instruments in the identification of the risk preference variables. These design characteristics are not directly correlated with the investment, income and endowment variables. A high level of significance of these design characteristics in the previously outlined regressions for risk preference parameters is a requirement for these instruments to be considered strong, see Tables 6-8 for the evidence. These variables do not only pick up noise, they also directly improve the estimates of the preference parameters.

The models were estimated as GLM with a log link which assumes $\ln(E[y/X]) = Xb$. This specification is suitable for regressions with nonnegative skewed dependent variables (Nichols 2010). It allows Y to be censored at zero and we had such observations for all four dependent variables. We estimated the following four alternative models for each of the four Y variables:

$$(14a) \quad Y_i = \exp\left(\mu^{DFT10} + \mu_{ce}\, r_i^{DFT10} + \mu_{ae}\, \alpha_i^{DFT10} + \mu_{be}\, \beta_i^{DFT10} + \mu_{aj}\, \xi_i^{DFT10} + \mu_{aej} E_{ji} + v_i^{DFT10} \right)$$

$$(14b) \quad Y_i = \exp\left(\mu^{DFT12} + \mu_{ce}\, r_i^{DFT12} + \mu_{ae}\, \alpha_i^{DFT12} + \mu_{be}\, \beta_i^{DFT12} + \mu_{aj}\, \xi_i^{DFT12} + \mu_{aej} E_{ji} + v_i^{DFT12} \right)$$

$$(14c) \quad Y_i = \exp\left(\mu^{CU10} + \mu_{ce}\, r_i^{CU10} + \mu_{ae}\, \overline{\alpha_i^{CU10}} + \mu_{be}\, \overline{\beta_i^{CU10}} + \mu_{aj}\, \overline{\xi_i^{CU10}} + \mu_{aej} E_{ji} + v_i^{CU10} \right)$$

$$(14d) \quad Y_i = \exp\left(\mu^{CU12} + \mu_{ce}\, \overline{r_i^{CU12}} + \mu_{ae}\, \overline{\alpha_i^{CU12}} + \mu_{be}\, \overline{\beta_i^{CU12}} + \mu_{aj}\, \overline{\xi_i^{CU12}} + \mu_{aej} E_{ji} + v_i^{CU12} \right)$$

As an additional robustness check, these models were run without the predicted noise parameters as well. Experimental enumerator fixed effects were included in all models. Standard errors were bootstrapped, resampling youth groups with 500 replications. The AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) measures of model fit were used to compare the different specifications without and with the predicted noise parameters and models with 10 CLs vs. models with 12 CLs. The significance and stability of the risk preference parameters were used as additional indicators of the predictive power. As additional robustness checks, random effects panel Poisson models were used for the number of durable assets variable and random effects panel Tobit models were used for the livestock endowment variable. In terms of parameter signs and significance, the key results were very robust to these alternative specifications, see Appendix Tables 11A and 12A for the results of the additional Poisson and Tobit specifications.

5. Results and discussion

5.1. Parameter distributions by CL inclusion and error specification

Structural models were estimated without and with the two last CLs (11 and 12) which included lower probability higher gain outcomes. This was done with the two alternative approaches to capture the contextual variation in CLs to allow heteroskedastic errors (CU of Wilcox 2011; 2015; and DFT of Busemeyer and Townsend 1992; 1993) where we have included controls for CL specific and individual characteristics. The models include experimental enumerator and district fixed effects as well. We first get an overview by looking at the predicted parameter distributions before we go to the tables with more detailed econometric results for the alternative models.

Figures 2a and 2b present the predicted parameter distributions with the DFT and CU specifications for models based on the first 10 CLs. We see that the distributions are quite similar and with the utility curvature (CPRRA parameter) distributed around risk neutrality. However, with the very limited variation and relatively small outcome levels in the 10 first series, we may

not expect to detect much utility curvature. The distribution of the Prelec 2 alpha parameter has the main peak around 0.5-0.6 and a smaller peak around 1 showing that the large majority of the respondents have the inverted S-shaped probability weighting function while a small group has close to linear probability weighting like for EUT. These results are quite in line with several studies utilizing large student samples. Bruhin et al. (2010) found that about 80% of their student samples from Switzerland and China had non-linear probability weighting while about 20% had more EUT like preferences with linear probability weighting. We see that the Prelec 2 beta parameters had the peak slightly above 1, indicating a weak pessimistic tendency for the majority.

In a cross-country study of risk preferences among university student samples that included Ethiopia, l'Haridon and Vieider (2019) found the Prelec 2 alpha parameter to be somewhat higher (just above 0.6) and the Prelec 2 beta parameter to be substantially lower (about 0.73), signaling more optimistic preferences. Vieider et al. (2018) assessed the risk preferences of a representative sample of 500 farmers in Ethiopia and found a high level of risk tolerance and high level of optimism also in their sample of farmers similar to that in the Ethiopian student sample. They had 7 CLs with probabilities varying from 0.05 to 0.95 and the good outcome was constant at 40 ETB and bad outcome at 0 across the CLs. The respondents were presented the full CL in each case and the CLs used 1 ETB intervals for the sure amounts. Our estimates for the low resource rural youth and young adult sample in Figures 2a and 2b for the 10 CLs are similar as those in the above studies for risk tolerance and probabilistic insensitivity but our sample is less optimistic than in their studies. The picture changes, however, when we add the two extra CLs with low probability high payouts (see Figures 3a and 3b).
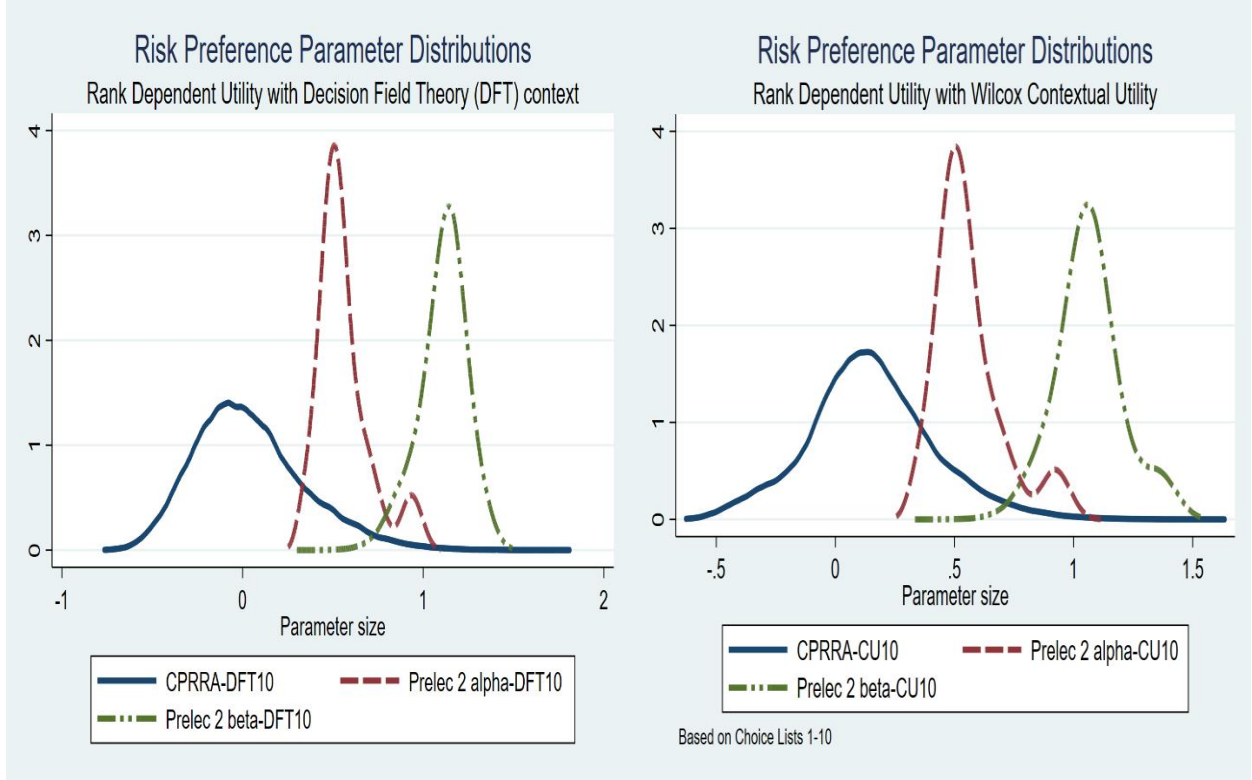
Figures 2a and 2b. Risk preference parameter distribution with 10 CLs and DFT versus Wilcox contextual utility (CU).

In Figures 3a and 3b the data for all 12 CLs have been included, otherwise they are estimated in the same way as in Figures 2a and 2b. We see a remarkable change in the distribution of the utility curvature (CPRRA) parameter with the peak now around 1, indicating substantial risk aversion. There is still a minority that is risk neutral or risk loving. This shows that the parameter estimates are very sensitive to the construction of CLs and the range of outcome values in these. The inclusion of the two CLs with low probability high gains gave more outcome variation that could pick up more utility curvature. But here is also another potential driver of the change related to the design of these two CLs. Figure 1b illustrated the cumulative distribution of switch points in the sample and we see that about 25% of the sample never switched for these two CLs, implying a censoring of the responses. This in combination with random choices could possibly explain biased estimates of risk aversion for a substantial share of the sample. This requires further examination.

When it comes to the Prelec 2 alpha parameter distribution, it was not much affected by the addition of the two extra CLs, while the Prelec 2 beta parameter distribution has shifted towards

the left in both Figures 3a and 3b, indicating more optimistic expectations while there is also a smaller group with more pessimistic expectations that is more pronounced in Figure 3b than in 3a.
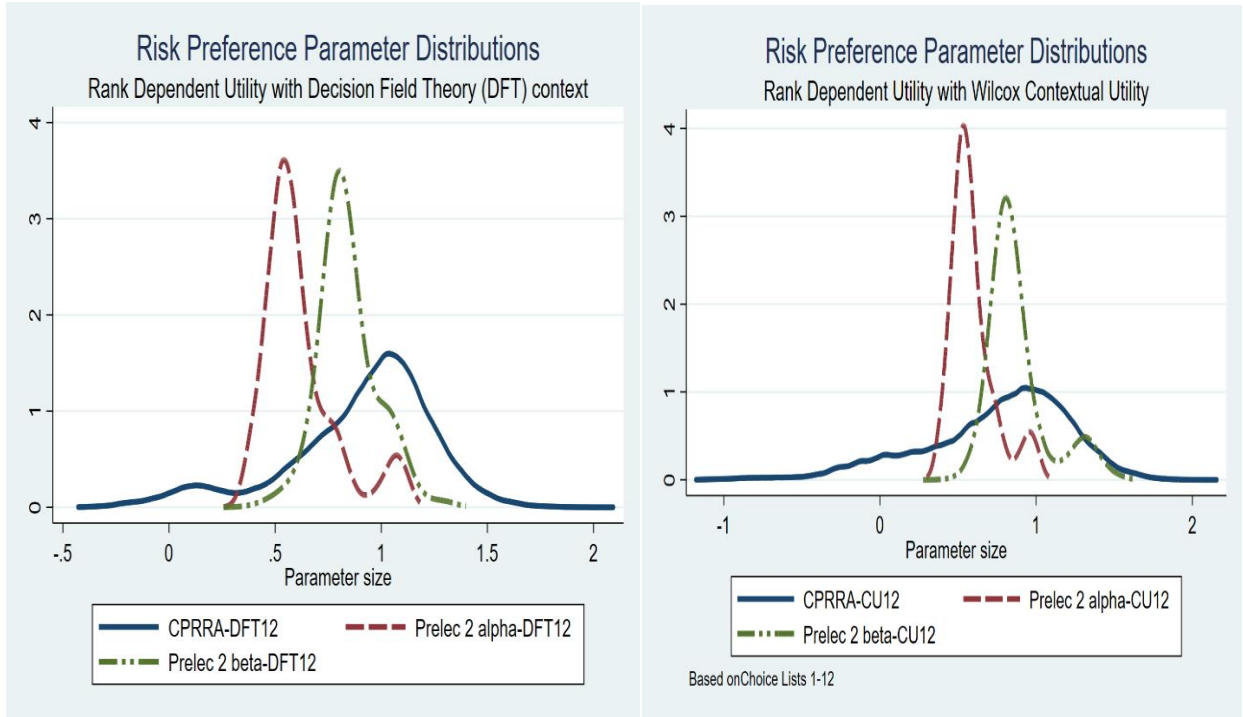
Figures 3a and 3b. Risk preference parameter distributions with 12 CLs and DFT vs. Wilcox contextual utility (CU).

Next, we assess the error distributions in the four model alternatives. Due to the differences in weighting in the CU and DFT models the errors in these are not directly comparable. We therefore compare pairwise the DFT10 and DFT12 error distributions and the CU10 vs. CU12 error distributions in Figures 4a and 4b. We suspect that the addition of the two last CLs has resulted in a wider spread in the errors, particularly given the extent of corner outcomes in Figure 1b. Figure 4a, with the DFT specifications, indicates that the spread may have increased but that the error size has been reduced for the majority of the sample. We see a weaker tendency of the same also in Figure 4b, with the CU specifications. We see that there is a share of the observations with errors that are two times or more the size of the peak of the distributions. Note that this is after we have controlled for potential order effects, the random starting point row number, the risk neutral task

number in the CL, the probability of bad outcome, the number of years of completed education, and experimental enumerator fixed effects in the error part of the structural models.
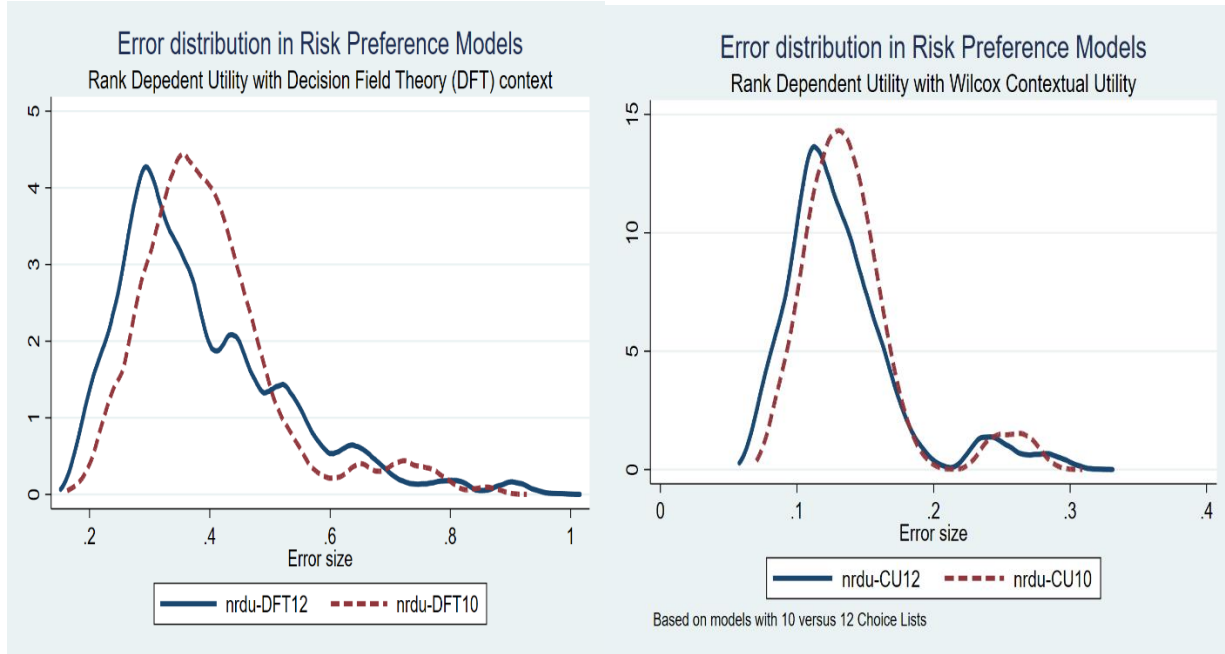
Figures 4a and 4b. Error distribution in RDU-DFT10/DFT12 versus RDU-CU10/CU12 models

5.2.Assessment of the stability and sensitivity of risk preference parameters to CL design and contextual econometric framing

We will now look more closely at the econometric model results. We start by inspecting the four models for the parsimonious specifications without any controls. These are presented in Table 3. We see the same sharp increase in the utility curvature parameter (CPRRA) from -0.19-0.03 to 0.80-0.98 after adding CL 11 and 12 with low probability high outcomes. However, we should keep in mind that these parameters may be biased due to order effects, random choices, the positions of the risk neutral task in each CL, and the random starting point in each CL. Furthermore, the Prelec 2 alpha parameter is increased from 0.45 to 0.57-0.60 with the addition of the two extra CLs while the Prelec 2 beta is reduced from 0.96-1.06 to 0.82-0.85 (more optimistic). Figure 5 illustrates the probability weighting functions for these four models.

Table 3. CU and DFT models with 12 and 10 Choice lists (without CL 11 and 12).

	(1)	(2)	(3)	(4)
	CU 12	CU 10	DFT 12	DFT 10
CPRRA	0.798***	-0.186**	0.983***	0.0345
	(0.059)	(0.067)	(0.035)	(0.071)
Prelec 2 alpha	0.574***	0.454***	0.596***	0.446***
	(0.008)	(0.007)	(0.008)	(0.007)
Prelec 2 beta	0.851***	1.055***	0.822***	0.961***
	(0.024)	(0.030)	(0.021)	(0.030)
CU/DFT Noise	0.152***	0.168***	0.444***	0.450***
	(0.004)	(0.004)	(0.010)	(0.012)
N	116577	98911	116577	98911

Note: Standard errors in parentheses, corrected for clustering at respondent level. * $p<0.05$, ** $p<0.01$, *** $p<0.001$
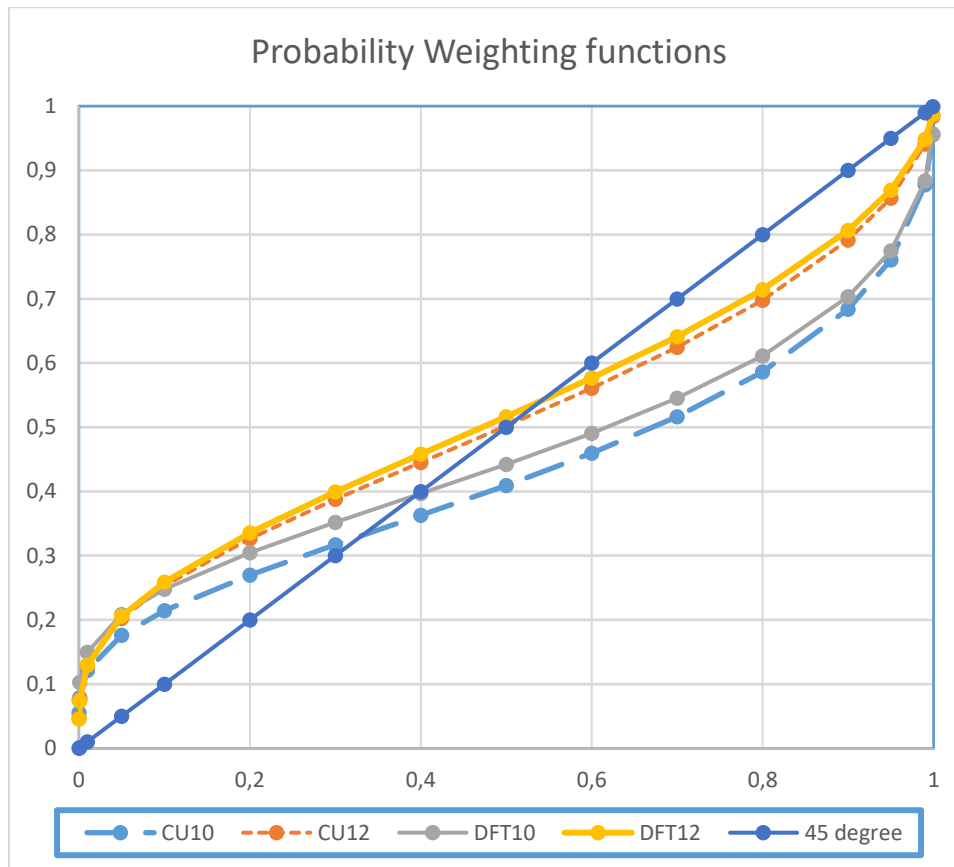
Figure 5. Probability weighting functions based on estimates in Table 3

Table 4 gives an overview of CL characteristics that have been included as controls. As an extra test for possible errors associated with cognitive ability, the number of years of completed education was included. We also included the probability of bad outcome as we put more emphasis on CLs with relatively low probability bad outcomes given our context. We therefore expect errors to be higher when the probability of bad outcome is higher. We expected that the majority of respondents would be in the risk averse zone for most of the CLs (assuming linear probability weighting), based on earlier studies of risk preferences in developing countries in general and Ethiopia in particular (Binswanger 1981; Wik et al. 2004; Yesuf et al. 2009) which found quite high levels of risk aversion. This is also due to the inverted S-shape of the probability weighting function, given that most of our CLs focused on the region with high probability of good outcomes (low probability of bad outcome). In the estimation we included the task row number where the linear-probability risk neutral row is located in the CL as a control variable to assess the potential bias associated with this position (Andersson et al. 2016), including the squared of this row number as well, to test for an eventual non-linear effect.

It is important to remember that we did not present the full CL to the respondents, just the binary choice associated with a randomly chosen starting row in the CL. The position of the risk neutral row is therefore only affecting the probability that it is inside or outside the region of the list determined by the random starting row and the decision made at that row.

Table 4. CL characteristics and education included as controls in the error specification

Variable	Obs	Mean	Std. Dev.	Min	Max
Page number (order of CL)	112,606	3.45	1.69	1	6
Starting point task number	112,606	0.54	1.84	0	10
Risk neutral task number	112,606	3.09	1.83	1	8
Prob. of Bad outcome	112,606	0.32	0.27	0.05	0.95
Education, years completed	112,606	5.39	3.96	0	17

Note: We made some robustness checks for the position of the risk neutral row relative to the starting point task number and for the starting point being in the risk-loving zone. These are included in Appendix 1.

Table 5 presents the results for the correlations between these controls and the noise parameter in the four models. First, we see a significant order effect for the errors as the errors decline with page number (indicating the sequence of the CLs). This may indicate a learning effect with

exposure to multiple CLs. We tested for a non-linear order effect as well but it was insignificant. Second, there is an effect of the position of the random starting point in the CLs. With a starting point further down in the list, the error is higher. This is an indication of random error in the first decision in each list. We will return to how this may have resulted in bias in the risk preference parameters below. Third, there is a significant non-linear effect of the position of the linear probability risk neutral row in the CL. The error is larger with a placement higher up on the list but the effect is declining as evidenced by the significant positive parameter for the quadratic variable. We interpret this as there being higher likelihood of random errors when the risk-neutral row is placed close to the top of the CLs. This is also in line with the findings of Andersson et al. (2016). Fourth, there is a tendency towards larger errors when the probability of bad outcome is higher even though this is only significant in one of the CU models, CU12, after the two high probability of bad outcome CLs were included. Adding those two CLs substantially increases the positive error correlation with this probability as we suspected. This also implies that we should expect higher accuracy for the predicted probability weighting function in the low probability of bad outcome region. Fifth, the education variable only became significant in the models with all 12 CLs but the signs in all models were negative. This is in line with our expectation that less education is associated with higher likelihood of error commitments. The lack of significance in two of the models indicates that the other sources of potential bias are more important than the general low and variable education level in our sample. Risk perceptions may be conceptually well understood also by individuals with limited education as they have grown up in this kind of risky environment. They may, however, be less used to assessing prospects with low probabilities of high gains and that may be why education is more important for the comprehension of the last two CLs. Andersson et al. (2018) suggested that age may be associated with cognitive ability model noise. We tested this as well but found no such significant correlation. This could be because our sample consists of young adults while age-related cognition problems tend to appear at older age.

Table 5. Model error and error sources in CU and DFT models

Noise	DFT10	DFT12	CU10	CU12
Page number	-0.0113***	-0.0067**	-0.0032***	-0.0038***
	(0.003)	(0.002)	(0.001)	(0.001)
Startp. Taskno	0.0059***	0.0091***	0.0021***	0.0029***
	(0.002)	(0.001)	(0.001)	(0.000)
Risk neutral Taskno	-0.103***	-0.0882***	-0.0322***	-0.0268***
	(0.006)	(0.006)	(0.003)	(0.002)
Risk neutral Taskno, squared	0.0100***	0.0092***	0.0035***	0.0030***
	(0.001)	(0.001)	(0.000)	(0.000)
Prob. Bad outcome	0.102***	0.268***	-0.016	0.0451***
	(0.022)	(0.018)	(0.012)	(0.008)
Education, years	-0.003	-0.0036*	-0.001	-0.0012*
	(0.002)	(0.001)	(0.001)	(0.000)
Enumerator FE	Yes	Yes	Yes	Yes
Constant	0.590***	0.470***	0.208***	0.170***
	(0.033)	(0.028)	(0.012)	(0.010)
N	95600	112606	95600	112606

Note: Standard errors in parentheses, corrected for clustering at respondent level. * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

Tables 6, 7 and 8 present the results for the structural models for each risk preference parameter. We have included the tests for order effects, random starting point row and the position of the linear probability risk neutral row in each CL, that we suspect may bias the parameter estimates. In addition, we included age, gender dummy, education, birth rank, and number of brothers as person characteristics and district fixed effects as controls for broader economic characteristics. We start by inspecting their correlations with the utility curvature parameter (Table 6).

Table 6 indicates a weak order effect, significant at 5% level in two out of the four model specifications and with positive signs in all four models. This tendency may also be a result of the learning that caused a reduction in the noise parameter as we saw in Table 5. It may indicate a weak tendency to underestimate the CPRRA parameter in the first CLs compared to later CLs. Second, the starting point task row number was significant to highly significant and with a positive sign in all models. This implies that a starting point further down in the list was associated with a

higher measured risk aversion parameter. This may indicate errors associated with the random starting point being near the top of the list and resulting in a movement to the top of the list in the second choice while the latent parameter is located below the starting point. The addition of the last two CLs reduced this upward starting point bias in risk aversion estimates, as it can be seen that the size of the parameter for starting point row is much reduced from the 10 CL to the 12 CL models. This is, however, more than compensated for with the much higher constant terms in the 12 CL models. Third, the linear probability risk neutral row number is significant in three of the four models but the signs are changing from positive in the 10 CL models to negative in the 12 CL models. We interpret this as an indication that placing the risk neutral row further down in the list in the first 10 CLs leads to a weak tendency towards higher estimates of risk aversion (the parameter was significant at 5% level only in one of these two models. The parameter for the risk neutral row number was, however, highly significant and with negative sign in both the 12 CL models. We interpret this as a strong indicator of bias towards overestimated risk aversion with these lists included if this bias is not corrected for. Many of the respondents are likely to be more willing to take risk in this probability region. The censoring of the safe amounts in these CLs lead to a bias in the CPRRA distribution, at least for a part of the sample (close to 25% appeared to be censored in Figure 1b). The inclusion of the risk neutral row number in these models may only partly have reduced this bias. Furthermore, we see very weak correlations between the CPRRA parameter and the observable individual characteristics and this was also the case for the district fixed effects. Finally, after the inclusion of the bias correction variables for CL design and the socioeconomic variables, we still see a large variation in the constant terms in the models, which may be an indicator of remaining uncorrected biases that vary across models.

Table 6. Utility curvature estimates (CPRRA) in CU and DFT models

CPRRA	DFT10	DFT12	CU10	CU12
Page number	0.0596*	0.019	0.039	0.0564*
	(0.027)	(0.026)	(0.025)	(0.022)
Startp. Taskno	0.0641***	0.0388**	0.0644***	0.0222*
	(0.013)	(0.012)	(0.011)	(0.009)
Risk neutral Taskno	0.0856*	-0.170***	0.070	-0.173***
	(0.038)	(0.026)	(0.045)	(0.021)
Age, years	0.001	0.005	0.003	0.005
	(0.007)	(0.006)	(0.008)	(0.005)
Male, dummy	0.054	-0.013	-0.021	-0.117
	(0.108)	(0.104)	(0.115)	(0.081)
Education, years	0.003	0.021	0.011	0.020
	(0.024)	(0.024)	(0.026)	(0.016)
Birth rank	0.011	0.026	0.016	0.024
	(0.028)	(0.022)	(0.030)	(0.021)
Number of brothers	0.050	0.034	0.068	0.0600*
	(0.035)	(0.026)	(0.039)	(0.027)
District Fixed Effects	Yes	Yes	Yes	Yes
Constant	-0.605*	1.123***	-0.583*	0.763**
	(0.297)	(0.296)	(0.296)	(0.271)

Note: Standard errors in parentheses, corrected for clustering at respondent level. * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Number of observations are the same as in Table 4.

The correlations between the Prelec 2 alpha parameter and the CL and socioeconomic observables are presented in Table 7. First, we see a significant order bias in the 12 CL models unlike in the 10 CL models, but the coefficients are positive in sign in all models. Learning may therefore have a tendency to reduce our measure of probabilistic insensitivity. Second, for the random starting point task row we have a strongly significant positive effect of higher task row number in the 10 CL models but not in the 12 CL models. This seems to be due to the same type of random error effect as we saw for the CPRRA parameter. A starting point high up in the list (small row number) may have increased the probability of random error and going towards the risk neutral/higher alpha

parameter region in the list and thereby leading to an upward bias in the Prelec 2 alpha parameter. It is intriguing that this significant bias has disappeared when adding the last two CLs. Third, the linear probability risk neutral row position was highly significant and positively associated with the Prelec 2 alpha parameter in all models. This may imply that having positioned the risk neutral row too high in the CLs may have resulted in a downward bias in the Prelec 2 alpha parameter in all models. This illustrates the sensitivity to the CL design even when the full CL is not presented to the respondents because the randomly sampled rows combined with random choice errors can lead to bias.

Table 7. Prelec 2 alpha estimates in DFT and CU models

Prelec 2 alpha	DFT10	DFT12	CU10	CU12
Page number	0.006	0.0094*	0.005	0.0077**
	(0.003)	(0.004)	(0.003)	(0.003)
Startp. Taskno	0.0069***	-0.003	0.0091***	0.000
	(0.002)	(0.002)	(0.002)	(0.001)
Risk neutral Taskno	0.0758***	0.0877***	0.0725***	0.0727***
	(0.006)	(0.004)	(0.006)	(0.005)
Age, years	-0.0017*	-0.0027**	-0.0023*	-0.0026***
	(0.001)	(0.001)	(0.001)	(0.001)
Male, dummy	0.006	0.010	0.014	0.022
	(0.015)	(0.016)	(0.016)	(0.013)
Education, years	-0.003	-0.004	-0.004	-0.003
	(0.002)	(0.003)	(0.002)	(0.002)
Birth rank	-0.007	-0.006	-0.008	-0.005
	(0.004)	(0.004)	(0.004)	(0.003)
Number of brothers	0.005	-0.002	0.004	-0.002
	(0.005)	(0.005)	(0.005)	(0.005)
District Fixed Effects	Yes	Yes	Yes	Yes
Constant	0.443***	0.470***	0.468***	0.488***
	(0.042)	(0.047)	(0.049)	(0.040)

Note: Standard errors in parentheses, corrected for clustering at respondent level. * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Number of observations are the same as in Table 4.

Among the respondent characteristics, it is only age that is significant and with a negative sign in all four models. Older age is therefore associated with a stronger inverted-S shape of the probability weighting function. Finally, we see that there is limited variation in the constant term for the parameter across the models. This may indicate that the bias correction terms have been able to correct much of the bias variation across models for the Prelec 2 alpha parameter.

Table 8 presents the correlations for the Prelec 2 beta parameter, which captures the elevation of the probability weighting function. First, we see no sign of order bias for this parameter but there

Table 8. Prelec 2 beta estimates in DFT and CU models

Prelec 2 beta	DFT10	DFT12	CU10	CU12
Page number	-0.016	0.009	-0.007	-0.008
	(0.013)	(0.013)	(0.012)	(0.009)
Startp. Taskno	-0.0239***	-0.0264***	-0.0215***	-0.0159***
	(0.005)	(0.005)	(0.005)	(0.003)
Risk neutral Taskno	-0.0428***	0.0449***	-0.0390**	0.0375***
	(0.011)	(0.008)	(0.014)	(0.007)
Age, years	-0.002	-0.004	-0.004	-0.004
	(0.004)	(0.003)	(0.005)	(0.003)
Male, dummy	-0.044	-0.004	0.003	0.045
	(0.058)	(0.059)	(0.065)	(0.040)
Education, years	-0.011	-0.016	-0.016	-0.015
	(0.013)	(0.013)	(0.015)	(0.008)
Birth rank	-0.014	-0.015	-0.017	-0.013
	(0.014)	(0.012)	(0.016)	(0.010)
Number of brothers	-0.018	-0.019	-0.027	-0.0286*
	(0.017)	(0.015)	(0.020)	(0.013)
District Fixed Effects	Yes	Yes	Yes	Yes
Constant	1.414***	0.889***	1.465***	0.967***
	(0.160)	(0.165)	(0.181)	(0.124)

Note: Standard errors in parentheses, corrected for clustering at respondent level. * p<0.05, ** p<0.01, *** p<0.001. Number of observations are the same as in Table 4.

is a highly significant effect of the position of the starting row, which is a sign of random choice error associated with the first decision in each CL. Such errors have in this case resulted in a

stronger downward bias (optimism bias) when the random starting point was further down in the CL. Or, it could have been associated with a pessimism bias if the starting point is high in the list (low row number). Third, we see that the position of the risk neutral task row is highly significant in all models but with opposite signs in the 10 CL versus the 12 CL models. As the risk neutral task row number is low in CLs 11 and 12, this seems to have created a downward (optimism bias) that has more than counteracted the pessimism bias associated with average low risk neutral task row number in the 10 CLs.

Finally, we see that none of the observable respondent characteristics were significantly correlated with the Prelec 2 beta parameter (except number of brothers in one model). It is perhaps surprising that the gender dummy variable was not significantly correlated with any of the risk preference parameters given that many studies have found women to on average to be more risk averse than men. E.g. Vieider et al. (2018) found such a gender difference in their sample of Ethiopian farmers. As only one third of our sample are women, there may have been a stronger selection for women such that the more risk tolerant women have joined these youth business groups. We are unable to test this with our data, however. In general, our study is in line with other studies that find limited correlation between observable individual characteristics and risk preference parameters (von Gaudecker et al. 2011; l'Haridon and Vieider 2019).

5.3. Correlations between risk preference parameters and investment, income and endowment variables: CL design and predictive power

We will now investigate the degree of correlation between the alternative predicted risk preference parameters with 10 versus 12 CLs for the CU and DFT contextual framing alternatives and a set of four investment, income and endowment variables. Many studies have found weak and unstable predictive power of risk preferences and recent studies have found that design errors may be one important source of these confusing and inconclusive findings (Chuang and Schechter 2015; Andersson et al. 2016; Vieider 2018). Our study is exploratory and investigates the predictive power after controlling for potential biases due to CL design, without and with the addition of the two extra low probability high return CLs, without and with the predicted errors, and with the alternative CU versus DFT contextual framing approaches (Wilcox 2011; 2015; Busemeyer and Townsend 1992; 1993).

Given the weak predictive power of elicited risk preference estimates in earlier studies and the insignificant correlation with observable respondent characteristics in the previous section, our H0 hypothesis is that "*Our predicted risk preference parameters are uncorrelated with investment, income and endowment variables of the respondents*". We test a set of alternative hypotheses against this H0 hypothesis. Our H1 hypothesis states that "*The predicted CPRRA parameter is negatively correlated with investment, income, and endowment variables*". Our H2 hypothesis states that "*The Prelec 2 alpha parameter is positively correlated with investment, income and endowment variables (a lower alpha indicates a more probabilistically risk averse individual who invests less and therefore makes less income)*". Our H3 hypothesis states that "*The Prelec 2 beta parameter is negatively correlated with the investment, income and endowment variables (more optimistic respondents invest more, have higher income and more endowments)*". Hypothesis H4 states that "*A higher noise parameter is associated with lower investment, income and endowment variables (more noise is associated with lower cognitive ability and lower cognitive ability is associated with poorer economic performance)*". On the 10 CL vs. 12 CL basis for the estimated parameters we test the following two hypotheses; H5 states that "*The addition of the two low probability high return outcome CLs improves the predictive power of the predicted risk preference parameters and particularly so for the utility curvature parameter in line with hypothesis H1*"; and H6 states that "*Addition of the extra CLs is associated with more measurement error and does not add to the predictive power of the predicted parameters*". For the CU versus DFT contextual framing approaches we do not have any hypothesis about which one is better.

We assess these hypotheses for four variables; a) the log value of investments last year (Table 9); b) the log value of individual income during last year (other than group income) (Table 10); c) the number of durable assets owned by the individual (Table 11); and d) the livestock endowment of the individual measured in Tropical livestock units (Table 12).

We see from Tables 9-12 that there is strong support for hypothesis H1. The hypothesis has strong support for all the dependent variables. The CPRRA parameter has a negative sign in all models and is significant in 31 of 32 models. Hypothesis H2 (Prelec 2 alpha parameter) has to be rejected. The parameter is negative in most cases and even significant and negative in many cases. Stronger probabilistic insensitivity (lower alpha) is associated with higher endowments and income, but was only significant in two of the eight investment models. Hypothesis H3 (Prelec 2 beta parameter)

is also strongly supported by the results. All models give negative parameter estimates and these are significant in 28 out of 32 models. More pessimistic preferences are associated with lower investments, income and endowment levels. Hypothesis H4 (noise as sign of limited cognitive ability and poor economic performance) has to be rejected. Surprisingly, the noise parameter has positive sign in all models. It is insignificant in two of the four investment models (Table 9) but is significant in all the remaining 12 models. This is a puzzling result that we do not have good explanations for. It is all the more puzzling as other studies using larger and broader samples with more heterogeneity have found stronger effects of cognitive ability and measurement error (Andersson et al. 2018; Dohmen et al. 2010; 2018).

Hypothesis H5 states that the two extra CLs add to the predictive power of the models cannot be rejected. In fact, the predictive power for the models with the 10 CLs was weak and it was substantially improved by adding the two extra CLs for all four outcome variables in Tables 9-12. Adding the noise parameter improved the fit of almost all model, with the exception of the 12 CL investment models in Table 9, as evidenced by the AIC and BIC fit characteristics. The coefficients on the CPRRA parameter became larger in absolute value and in many cases more significant in the models with 12 CLs compared to the matched 10 CL models. The coefficients on the noise parameters were lower in the 12 CL versus 10 CL models, and the parameters on the predicted risk preference parameters were more stable in the 12 CL models than the 10 CL models when comparing the models without and with the predicted noise parameters. This is indicating that the measurement error problem did not get worse with the addition of the two extra CLs, rather the opposite. The additional controls for potential design weaknesses appear to have worked well. The identification strategy using the randomization of order and starting point variables as instruments appears also to have worked well, and particularly so in the 12 CL models. We therefore have to reject hypothesis H6. Finally, when we compare the Wilcox Contextual Utility (CU) and the DFT models, both perform well in the 12 CL cases and give similar results.

Table 9. Correlations between investment and predicted risk preference parameters from models with 10 CL vs. 12 CL with DFT and CU specification, models without and with predicted noise parameters.

loginvest	(1) DFT10	(2) DFT10+e	(3) DFT12	(4) DFT12+e	loginvest	(5) CU10	(6) CU10+e	(7) CU12	(8) CU12+e
Rrdu-DFT10	-0.520***	-0.619***			Rrdu-CU10	-0.613***	-0.679***		
	(0.127)	(0.132)				(0.122)	(0.122)		
Ardu-DFT10	-0.096	-0.452			Ardu-CU10	0.119	-0.131		
	(0.255)	(0.255)				(0.219)	(0.225)		
Brdu-DFT10	-0.572**	-0.848***			Brdu-CU10	-0.643***	-0.884***		
	(0.203)	(0.217)				(0.174)	(0.191)		
Nrdu-DFT10		2.477**			Nrdu-CU10		6.919**		
		(0.910)					(2.591)		
Rrdu-DFT12			-1.428***	-1.429***	Rrdu-CU12			-1.012***	-1.015***
			(0.232)	(0.242)				(0.171)	(0.171)
Ardu-DFT12			-0.884***	-0.918***	Ardu-CU12			-0.366	-0.433
			(0.214)	(0.240)				(0.246)	(0.252)
Brdu-DFT12			-1.933***	-1.937***	Brdu-CU12			-1.625***	-1.640***
			(0.334)	(0.344)				(0.312)	(0.312)
Nrdu-DFT12				0.184	Nrdu-CU12				1.530
				(0.594)					(1.881)
Constant	2.717***	2.282***	5.451***	5.408***		2.678***	2.169***	4.397***	4.264***
	(0.304)	(0.339)	(0.553)	(0.561)		(0.235)	(0.324)	(0.481)	(0.500)
AIC	4275.9	4270.6	4236.4	4238.3		4258	4251.6	4227.6	4229
BIC	4348.6	4348.2	4309.2	4316.0		4330.8	4329.2	4300.4	4306.6
N	945	945	945	945		945	945	945	945

Note: Glm log link models. Dependent variable: Log of individual investment last year. Bootstrapped standard errors in parentheses, resampling youth groups, 500 replications. * p<0.05, ** p<0.01, *** p<0.001. Rrdu=CPRRA parameter, Ardu=Prelec 2 alpha parameter, Brdu=Prelec 2 beta parameter, Nrdu=Noise parameter. All models with experimental enumerator fixed effects. Model fit characteristics: AIC (Akaike Information Criterion) = -2*ln(likelihood) + 2*k. BIC (Bayesian Information Criterion) = -2*ln(likelihood) + ln(N)*k.

Table 10. Correlations between income and predicted risk preference parameters from models with 10 CL vs. 12 CL with DFT and CU specification, models without and with predicted noise parameters. All models with experimental enumerator fixed effects.

loindinc	(1) DFT10	(2) DFT10+e	(3) DFT12	(4) DFT12+e	loindinc	(5) CU10	(6) CU10+e	(7) CU12	(8) CU12+e
Rrdu-DFT10	-0.323	-0.781***			Rrdu-CU10	-0.538**	-0.814***		
	(0.221)	(0.185)				(0.192)	(0.162)		
Ardu-DFT10	-0.677	-2.229***			Ardu-CU10	-0.503	-1.526***		
	(0.350)	(0.341)				(0.390)	(0.328)		
Brdu-DFT10	-0.387	-1.691***			Brdu-CU10	-0.565	-1.646***		
	(0.352)	(0.340)				(0.294)	(0.269)		
Nrdu-DFT10		12.00***			Nrdu-CU10		33.59***		
		(1.543)					(3.853)		
Rrdu-DFT12			-2.387***	-2.325***	Rrdu-CU12			-1.724***	-1.678***
			(0.299)	(0.317)				(0.226)	(0.198)
Ardu-DFT12			-2.340***	-3.314***	Ardu-CU12			-1.948***	-2.651***
			(0.399)	(0.459)				(0.438)	(0.446)
Brdu-DFT12			-3.388***	-3.376***	Brdu-CU12			-3.016***	-3.050***
			(0.416)	(0.457)				(0.408)	(0.357)
Nrdu-DFT12				6.299***	Nrdu-CU12				21.52***
				(1.120)					(3.342)
Constant	2.990***	0.725	8.595***	6.822***		3.103***	0.407	7.250***	5.005***
	(0.480)	(0.554)	(0.749)	(0.863)		(0.391)	(0.498)	(0.663)	(0.657)
AIC	4993.4	4925.6	4920.1	4880.8		4984.4	4904.7	4926.3	4880
BIC	5066.1	5003.2	4992.8	4958.5		5057.2	4982.3	4999.1	4957.6
N	945	945	945	945		945	945	945	945

Note: Glm log link models. Dependent variable: loindinc=Log of other individual income last year. Bootstrapped standard errors in parentheses, resampling youth groups, 500 replications. * p<0.05, ** p<0.01, *** p<0.001. Rrdu=CPRRA parameter, Ardu=Prelec 2 alpha parameter, Brdu=Prelec 2 beta parameter, Nrdu=Noise parameter. All models with experimental enumerator fixed effects.

Table 11. Correlations between number of durable assets and predicted risk preference parameters from models with 10 CL vs. 12 CL with DFT and CU specification, linear panel models with enumerator fixed effects, without and with predicted noise parameters.

durassetno	(1) DFT10	(2) DFT10+e	(3) DFT12	(4) DFT12+e	durassetno	(5) CU10	(6) CU10+e	(7) CU12	(8) CU12+e
Rrdu-DFT10	-2.296**	-3.319***			Rrdu-CU10	-2.838***	-3.254***		
	(0.719)	(0.399)				(0.630)	(0.358)		
Ardu-DFT10	-1.231	-5.624***			Ardu-CU10	-1.237	-3.553***		
	(1.022)	(0.951)				(0.900)	(0.601)		
Brdu-DFT10	-1.068	-4.792***			Brdu-CU10	-2.933**	-5.508***		
	(1.029)	(0.712)				(0.987)	(0.571)		
Nrdu-DFT10		34.62***			Nrdu-CU10		99.68***		
		(3.739)					(8.164)		
Rrdu-DFT12			-7.548***	-6.579***	Rrdu-CU12			-5.664***	-4.746***
			(0.660)	(0.637)				(0.536)	(0.499)
Ardu-DFT12			-5.847***	-8.202***	Ardu-CU12			-5.518***	-6.692***
			(0.763)	(0.795)				(0.847)	(0.870)
Brdu-DFT12			-10.50***	-9.241***	Brdu-CU12			-9.975***	-8.529***
			(0.981)	(0.920)				(0.994)	(0.899)
Nrdu-DFT12				17.80***	Nrdu-CU12				64.49***
				(2.313)					(6.640)
Constant	2.321	-4.312***	19.40***	12.37***		4.321***	-4.772***	16.43***	7.203***
	(1.422)	(1.211)	(1.645)	(1.603)		(1.295)	(1.075)	(1.474)	(1.463)
AIC	3289.9	3113.5	3126.5	3022		3269.4	3040.3	3134	2998.4
BIC	3362.7	3191.2	3199.3	3099.6		3342.1	3118	3206.8	3076
N	945	945	945	945		945	945	945	945

Note: Glm log link models. Dependent variable: durassetno=Number of durable assets. Bootstrapped standard errors in parentheses, resampling youth groups, 500 replications. * p<0.05, ** p<0.01, *** p<0.001. Rrdu=CPRRA parameter, Ardu=Prelec 2 alpha parameter, Brdu=Prelec 2 beta parameter, Nrdu=Noise parameter. All models with experimental enumerator fixed effects.

Table 12. Correlations between livestock endowment and predicted risk preference parameters from models with 10 CL vs. 12 CL with DFT and CU specification, glm log link models with enumerator fixed effects, without and with predicted noise parameters.

TLU	(1) DFT10	(2) DFT10+e	(3) DFT12	(4) DFT12+e	TLU	(5) CU10	(6) CU10+e	(7) CU12	(8) CU12+e
Rrdu-DFT10	-1.820*	-2.862***			Rrdu-CU10	-2.862***	-3.337***		
	(0.869)	(0.623)				(0.709)	(0.650)		
Ardu-DFT10	0.753	-3.169			Ardu-CU10	1.035	-1.350		
	(2.529)	(3.520)				(2.079)	(2.758)		
Brdu-DFT10	-1.267	-5.471**			Brdu-CU10	-2.917**	-5.767***		
	(1.124)	(1.845)				(0.989)	(1.633)		
Nrdu-DFT10		38.06***			Nrdu-CU10		107.8***		
		(6.978)					(17.850)		
Rrdu-DFT12			-9.677***	-8.395***	Rrdu-CU12			-7.181***	-6.339***
			(1.392)	(1.309)				(1.052)	(0.943)
Ardu-DFT12			-4.422***	-6.735***	Ardu-CU12			-2.972*	-4.002***
			(1.165)	(0.968)				(1.439)	(1.216)
Brdu-DFT12			-13.55***	-11.82***	Brdu-CU12			-12.33***	-10.91***
			(1.910)	(1.795)				(1.746)	(1.552)
Nrdu-DFT12				18.86***	Nrdu-CU12				68.52***
				(3.522)					(12.030)
Constant	1.504	-6.182*	23.07***	14.94***		3.184	-6.636*	18.18***	8.462***
	(1.882)	(2.915)	(2.703)	(2.337)		(1.740)	(2.634)	(2.122)	(2.029)
AIC	3606	3498.5	3441.3	3375.1	AIC	3579.7	3439.9	3450.2	3358.4
BIC	3678.8	3576.1	3514.1	3452.7	BIC	3652.5	3517.5	3523	3436
N	945	945	945	945	N	945	945	945	945

Note: Dependent variable: TLU=Livestock endowment (Tropical livestock units Bootstrapped standard errors in parentheses, resampling youth groups, 500 replications. * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Rrdu=CPRRA parameter, Ardu=Prelec 2 alpha parameter, Brdu=Prelec 2 beta parameter, Nrdu=Noise parameter. See Table 12A in the appendix showing the results from alternative panel Tobit models with enumerator random effects.

Finally, we wanted to assess the robustness of the predictions to the inclusion of additional controls and the correlations between the investment, endowment and income variables. This assessed in Table 13 for the DFT models with 12 CLs. Model (1) is the last model in Table 9, while we have included the durable assets and livestock endowment variables from Tables 11 and 12 as RHS variables in model (2) in Table 13. We see that the coefficients for risk preference parameters are reduced but retain the same signs and the coefficient for Prelec 2 alpha becomes insignificant. The CPRRA and Prelec 2 beta results remained signifant also after the inclusion of income in model (3) and other respondent charactersiticst in model (4). Figure 6 shows the predictived marginal effects for the three risk preference parameters for models (1) -(3).

Table 13. DFT-CL12 Investment models with additional controls

	(1) DFT_CL12a	(2) DFT_CL12b	(3) DFT_CL12c	(4) DFT_CL12d
Rrdu-DFT12	-1.429***	-0.963***	-0.909***	-1.467**
	(0.253)	(0.206)	(0.217)	(0.491)
Ardu-DFT12	-0.918***	-0.299	-0.219	-1.429
	(0.260)	(0.241)	(0.232)	(0.811)
Brdu-DFT12	-1.937***	-1.285***	-1.205***	-2.102**
	(0.360)	(0.305)	(0.315)	(0.723)
Nrdu-DFT12	0.184	-1.041	-1.186*	-0.716
	(0.619)	(0.578)	(0.588)	(1.419)
Tropical livestock units		0.003	0.003	0.001
		(0.008)	(0.008)	(0.007)
Durable assets, number		0.0331***	0.0311***	0.0349***
		(0.008)	(0.008)	(0.008)
Log individual income			0.004	0.003
			(0.004)	(0.004)
Male, dummy				0.0898***
				(0.022)
Age, years				-0.00654*
				(0.003)
Education, years completed				-0.008
				(0.006)
Enumerator fixed effects	Yes	Yes	Yes	Yes
Constant	5.408***	4.455***	4.308***	6.328***
	(0.586)	(0.498)	(0.528)	(1.343)
N	945	945	945	945

Note: Dependent variable: Log of individual investment last year. Bootstrapped standard errors in parentheses, resampling youth business groups, 500 replications. * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Rrdu=CPRRA parameter, Ardu=Prelec 2 alpha parameter, Brdu=Prelec 2 beta parameter, Nrdu=Noise parameter. All models with experimental enumerator fixed effects.

Figure 6. Predicted marginal responses to risk preference parameters for DFT models with 12 CL with additional controls: a) error & enumerator FE, b) & endowments, c) & income.

5.4. Does loss aversion make a difference?

We did not measure loss aversion in the same comprehensive way as the other risk preference measures. However, we included one CL to get a measure of loss aversion, see Appendix 3. This CL allowed choice between two risky prospects at each low in the list with 50% chance of negative outcomes, and 50% chance of positive outcomes. The two risky prospects on each row differ in

size of negative outcomes and expected value. A switch point further down in the list (higher row number) indicates higher loss aversion (a preference of prospect A with smallest loss even though expected return is lower than for prospect B. We used the row number just below the switch point as a measure of loss aversion rank. Also for this CL the rapid elicitation method was used. We ran a separate regression to assess whether the loss aversion rank was correlated with the predicted risk preference and noise parameters, the random starting point in the loss aversion CL and observable respondent characteristics.

In Table 14 we included to row number for random starting row in the loss aversion CL. This CL was not included for about 10% of the sample. In Table 14 we have included models to first inspect the effect of this loss of observations although we did not expect it to affect the results significantly. Then we have included the loss aversion variable with the control for potential starting point bias in the third model of each of the DFT12 and CU12 model versions, and, finally, we have included additional respondent characteristics as controls.

Table 14 shows that the key parameters only changed slightly for the reduced sample. The key finding is that loss aversion was significant (at 1% level) and negatively correlated with investment. Higher levels of loss aversion seem to be associated with less investment. This result is robust to the inclusion of additional controls and is similar in the DFT and CU specifications. There was no significant effect of the random starting point, a source of measurement error in loss aversion. This variable was significant when loss aversion rank was regressed on it, risk preferences, noise and other variables, see Appendix Table A14. The Appendix Table also shows that loss aversion rank was not significantly correlated with the other risk preference parameters or noise in the structural models. Men were significantly less loss averse than women, while we recall that we did not find such a significant gender difference for the other risk preference parameters in the structural models. However, we see in Table 14 that men also invested significantly more than women after we have controlled for loss aversion, other risk preferences, measurement error and other observable respondent controls.

Table 14. The additional value of loss aversion as a predictor of investment behavior

	DFT12-1	DFT12-2	DFT12-3	DFT12-4		CU12-1	CU12-2	CU12-3	CU12-4
Rrdu-DFT12	-1.429***	-1.375***	-1.354***	-1.478*	Rrdu-CU12	-1.015***	-0.993***	-0.977***	-0.809*
	(0.246)	(0.265)	(0.253)	(0.666)		(0.166)	(0.184)	(0.172)	(0.362)
Ardu-DFT12	-0.918***	-1.121***	-1.108***	-1.662	Ardu-CU12	-0.433	-0.655*	-0.670*	-0.757
	(0.243)	(0.308)	(0.288)	(1.063)		(0.246)	(0.297)	(0.288)	(0.752)
Brdu-DFT12	-1.937***	-1.906***	-1.874***	-2.103*	Brdu-CU12	-1.640***	-1.637***	-1.615***	-1.347*
	(0.354)	(0.380)	(0.364)	(0.955)		(0.299)	(0.338)	(0.315)	(0.649)
Nrdu-DFT12	0.184	0.500	0.456	0.850	Nrdu-CU12	1.530	1.779	1.754	6.918
	(0.608)	(0.664)	(0.617)	(2.050)		(1.917)	(1.965)	(2.027)	(13.660)
Loss aversion rank			-0.0159**	-0.0132**				-0.0154**	-0.0131**
			(0.005)	(0.005)				(0.005)	(0.005)
LA stp			0.000	0.001				0.000	0.001
			(0.004)	(0.004)				(0.004)	(0.004)
Tropical livestock units				-0.003					-0.003
				(0.008)					(0.008)
Durable assets, number				0.0357***					0.0348***
				(0.009)					(0.009)
Log individual income				0.002					0.002
				(0.004)					(0.004)
Male, dummy				0.0801**					0.0750*
				(0.028)					(0.030)
Age, years				-0.006					-0.003
				(0.004)					(0.003)
Education, years completed				-0.002					0.008
				(0.008)					(0.017)
Enumerator FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	5.408***	5.336***	5.389***	5.949***		4.264***	4.338***	4.409***	3.377*
	(0.580)	(0.675)	(0.639)	(1.752)		(0.460)	(0.557)	(0.541)	(1.668)
AIC	4238.3	3792.7	3786.5	3765.3		4229.0	3786.7	3781.2	3764.4
BIC	4316.0	3868.5	3871.8	3879.0		4306.6	3862.5	3866.5	3878.0

N	945	843	843	843	945	843	843	843

Note: Dependent variable: Log of individual investment last year. Bootstrapped standard errors in parentheses, resampling youth business groups, 500 replications. * p<0.05, ** p<0.01, *** p<0.001. Rrdu=CPRRA parameter, Ardu=Prelec 2 alpha parameter, Brdu=Prelec 2 beta parameter, Nrdu=Noise parameter. All models with experimental enumerator fixed effects. Loss aversion is measured as rank by the row number below the switch point in the CL. LA stp is a control for the randomized starting point row in the CL testing for starting point bias. Models DFT12-1 vs DFT12-2 and CU12-1 vs CU12-2 assess the effect of attrition related to loss of observations when including the loss aversion variable (we do not have the loss aversion CL included for about 10% of the sample).

6. Conclusions

Our study aimed to contribute to the literature on elicitation of risk preferences in a developing country setting for respondents with limited education and that are searching for better livelihoods in a risky environment. We elicited relatively comprehensive measures of risk preferences based on rank dependent utility and a Prelec 2-parameter probability weighting function based on a Multiple Choice List approach and a new approach to rapid pairwise elicitation of certainty equivalents of risky prospects.

In line with Andersson et al. (2016) we found that the estimated risk preference parameters were sensitive to random errors and their interactions with the design of Choice Lists, and that these potentially can lead to substantial bias in the predicted parameters. The obvious solution may not be to place the linear probability risk neutral task in the middle of the list, however, when this is not a good representation of the average respondent. A two-stage approach may be recommended for the placement of the risk neutral row in each CL. A pilot study may help to identify a proxy probability weighting function. An identified inverted-S shaped probability weighting function implies that the middle row in a list should be in the risk-seeking area for low probability high outcome CLs while the middle row should be in the risk averse area for the low probability low outcome CLs. Still this kind of adjustment does not eliminate potential biases for respondents with tail-end preferences when random choice can cause censoring of the error distribution. Additional controls for CL design may help to control for such biases.

We found that our approach with random starting point and quickly narrowing in on the switch point has several advantages. It reduces the number of questions needed to find the switch point in each CL. It identifies only one switch point per series. It should be less prone to bias towards the middle of the list or anchoring towards the top or bottom of the lists. The rapid approach still has some problems. It is sensitive to random choice bias associated with the randomized staring point in each CL. It is also sensitive to the CL format such as location of risk-neutral row which affects the probability whether the starting point falls above or below this point. Still, the time saved for each CL can be used to include more CLs within the same time and budget frame, with extra power to control for CL design bias and starting point bias during estimation. We therefore think there is a net gain from using this approach compared to going through each CL from the top to the bottom

or bottom-up. This may be particularly helpful in environments with computer-illiterate respondents where experimental enumerators are necessary.

Ten of our CLs were designed to map the probability weighting function in the 5-50% probability of bad outcome region as these were the types of risks we though were most common in the environment of our respondents. However, we added two CLs with low probability high outcome risky prospects to assess whether this could help to better map the probability weighting function and separately identify the curvature of the utility function. We realized afterwards that we had not allowed sufficient variation in the risk loving zone for the safe amounts in these lists and this resulted in a share of respondents with no switch point. Random choice could therefore potentially contribute to more bias and weaker predictive power with the inclusion of these lists. The addition of these lists resulted in a sharp increase in the estimated utility curvature even after we attempted to control for CL design bias, which may not have been completely eliminated. However, the addition of these two CLs substantially improved the predictive power of the risk preference parameters in terms of their correlations with investment, income and endowment variables that also became less sensitive to the inclusion of the predicted noise parameters from the models.

We did not find that the noise parameter was strongly correlated with education and poor economic performance, rather the contrary. This indicates that our CL design and elicitation approach works quite well also for respondents with limited education. We think our findings and the lessons from our design errors provide valuable insights for the future design of more optimal instruments. We suspect that the large variation in estimates of risk preferences and the weak predictive power that has been found in many other studies can be due to similar weaknesses in design (Chuang and Schechter 2015; Dohmen et al. 2018; Schildberg-Hörisch 2018; Vieider 2018; Zhou and Hey 2018). More research is needed to further fine-tune experimental designs to different contexts in ways that still facilitate assessment of external validity. There are definitely still several devils hidden in the details.

References

Abdellaoui, M., Driouchi, A., & L'Haridon, O. (2011). Risk aversion elicitation: Reconciling tractability and bias minimization. *Theory and Decision*, 71, 63-80.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2006). Elicitation using multiple price list formats. *Experimental Economics*, 9(4), 383-405.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2008). Eliciting risk and time preferences. *Econometrica*, 76(3), 583-618.

Andersson, O., Holm, H. J., Tyran, J.-R., and Wengström, E. (2016). Risk Aversion relates to Cognitive Ability: Preferences or Noise? *Journal of the European Economic Association* 14(5), 1129–54.

Andersson, O., Holm, H. J., Tyran, J. R., & Wengström, E. (2018). Robust Inference in Risk Elicitation Tasks. *Available at SSRN 3256637*.

Apesteguia, J., and Ballester, M. A. (2018). Monotone Stochastic Choice Models: The Case of Risk and Time Preferences. *Journal of Political Economy*, 126 (1), 74–106.

Arrow, K. J. (1965). *Aspects of the theory of risk-bearing*. Yrjö Jahnssonin Säätiö lectures, Helsinki.

Benjamin, D. J., Brown, S. A., & Shapiro, J. M. (2013). Who is 'behavioral'? Cognitive ability and anomalous preferences. *Journal of the European Economic Association*, 11(6), 1231-1255.

Binswanger, H. P. (1981). Attitudes toward risk: Theoretical implications of an experiment in rural India. *Economic Journal*, 91(364), 867-890.

Bosch-Domènech, A., & Silvestre, J. (2013). Measuring risk aversion with lists: A new bias. *Theory and Decision*, 75(4), 465–496.

Bruhin, A., Fehr-Duda, H., & Epper, T. (2010). Risk and rationality: Uncovering heterogeneity in probability distortion. *Econometrica*, 78(4), 1375-1412.

Bruner, D. M. (2009). Changing the probability versus changing the reward. *Experimental Economics*, 12(4), 367-385.

Burks, S. V., Carpenter, J. P., Goette, L., & Rustichini, A. (2009). Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Sciences*, 106(19), 7745-7750.

Busemeyer, J. R., & Townsend, J. T. (1992). Fundamental derivations from decision field theory. *Mathematical Social Sciences*, 23(3), 255-282.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*(3), 432.

Camerer, C.F. (1995): Individual Decision Making. In *Handbook of Experimental Economics*, Vol. 2, ed. by A. R. J. Kagel. Princeton: Princeton University Press.

Callen, M., Isaqzadeh, M., Long, J. D., & Sprenger, C. (2014). Violence and risk preference: Experimental evidence from Afghanistan. *American Economic Review*, 104(1), 123-48.

Cerreia-Vioglio, S., Dillenberger, D., & Ortoleva, P. (2015). Cautious expected utility and the certainty effect. *Econometrica*, 83(2), 693-728.

Charness, G., & Viceisza, A. (2016). Three Risk-elicitation Methods in the Field: Evidence from Rural Senegal. *Review of Behavioral Economics*, 3(2), 145-171.

Chuang, Y., & Schechter, L. (2015). Stability of experimental and survey measures of risk, time, and social preferences: A review and some new results. *Journal of Development Economics*, 117, 151-170.

Conte, A. & Hey, J. D. (2018). Rehabilitating the Random Utility Model. A comment on Apesteguia and Ballester (2018). Discussion Paper 18/12. Department of Economics and Related Studies, University of York.

Csermely, T., & Rabas, A. (2016). How to reveal people's preferences: Comparing time consistency and predictive power of multiple price list risk elicitation methods. *Journal of Risk and Uncertainty*, 53(2-3), 107-136.

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.

Dohmen T., Falk, A., Huffman, D. and Sunde, U., (2010). Are Risk Aversion and Impatience Related to Cognitive Ability? *The American Economic Review*, 100(3): 1238-1260.

Drichoutis, A., & Lusk, J. (2012). Risk preference elicitation without the confounding effect of probability weighting. Working Paper, Munich, Germany.

Drichoutis, A. C., & Lusk, J. L. (2016). What can multiple price lists really tell us about risk preferences?. *Journal of Risk and Uncertainty*, 53(2-3), 89-106.

Freeman, D. J., Halevy, Y., & Kneeland, T. (2019). Eliciting risk preferences using choice lists. *Quantitative Economics*, 10(1), 217-237.

Friedman, D., Habib, S., James, D., & Crockett, S. (2019). *Varieties of risk elicitation*. WZB Discussion Paper. Version dated March 26, 2019.

Habib, S., Friedman, D., Crockett, S., & James, D. (2017). Payoff and presentation modulation of elicited risk preferences in MPLs. *Journal of the Economic Science Association*, *3*(2), 183-194.

Harrison, G. W., Johnson, E., McInnes, M. M., & Rutström, E. E. (2005). Risk aversion and incentive effects: Comment. *American Economic Review*, 95(3), 897-901.

Hey, J. D., Morone, A., & Schmidt, U. (2009). Noise and bias in eliciting preferences. *Journal of Risk and Uncertainty*, 39(3), 213-235.

Lévy-Garboua, L., Maafi, H., Masclet, D., & Terracol, A. (2012). Risk aversion and framing effects. *Experimental Economics*, 15(1), 128–144.

L'Haridon, O., Vieider, F. (2019) All over the map: A Worldwide Comparison of Risk Preferences. *Quantitative Economics*, 10, 185-215.

Loomes, G. & Pogrebna, G. (2014). Measuring Individual Risk Attitudes when Preferences are Imprecise. *The Economic Journal*, 124(576), 569–593. https://doi.org/10.1111/ecoj.12143

Nichols, A. (2010). Regression for nonnegative skewed dependent variables. https://www.stata.com/meeting/boston10/boston10_nichols.pdf

Oechssler, J., Roider, A., & Schmitz, P. W. (2009). Cognitive abilities and behavioral biases. *Journal of Economic Behavior & Organization*, 72(1), 147-152.

Poulton, E. C., & Poulton, S. (1989). *Bias in quantifying judgements*. Taylor & Francis.

Pratt, J. W. (1964) Risk aversion in the Small and in the Large. *Econometrica,* 32, 122-136.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization* 3(4) 323-343.

Rabin, M. (2000). Risk Aversion and Expected-Utility Theory: A Calibration Theorem. *Econometrica*, 68(5), 1281-1292.

Schildberg-Hörisch, H. (2018). Are risk preferences stable? *Journal of Economic Perspectives*, 32(2), 135-54.

Tanaka, T., Camerer, C. F., & Nguyen, Q. (2010). Risk and time preferences: Linking experimental and household survey data from Vietnam. *American Economic Review*, 100(1), 557-571.

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

Vieider, F. M. (2018). Violence and risk preference: experimental evidence from Afghanistan: comment. *American Economic Review*, 108(8), 2366-2382.

Vieider, F. M., Beyene, A., Bluffstone, R., Dissanayake, S., Gebreegziabher, Z., Martinsson, P., & Mekonnen, A. (2018). Measuring risk preferences in rural Ethiopia. *Economic Development and Cultural Change*, 66(3), 417-446.

Von Gaudecker, H. M., Van Soest, A., & Wengstrom, E. (2011). Heterogeneity in risky choice behavior in a broad population. *American Economic Review*, 101(2), 664-694.

Weber, E. U., Blais, A. R., & Betz, N. E. (2002). A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making*, 15(4), 263-290.

Wik, M., Aragie Kebede, T., Bergland, O., & Holden, S. T. (2004). On the measurement of risk aversion from experimental data. *Applied Economics*, 36(21), 2443-2451.

Wilcox, N. T. (2011). 'Stochastically more risk averse:'A contextual theory of stochastic discrete choice under risk. *Journal of Econometrics*, *162*(1), 89-104.

Wilcox, N. T. (2015). *Error and Generalization in Discrete Choice Under Risk* (No. 15-11).

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.

Zhou, W., & Hey, J. (2018). Context matters. *Experimental economics*, 21(4), 723-756.

Appendix 1. Additional tests for CL design biases: Starting point bias and position of risk neutral row

Table A1. Checking for biases related to choice list design, ordering of choice lists, and location of random starting point

Models with all 12 CLs	(1)	(2)	(3)	(4)	(5)
CPRRA					
Page number	-0.0546*	-0.0562**	-0.0519*	-0.0539*	-0.0579**
	(-2.11)	(-2.71)	(-2.04)	(-2.29)	(-2.92)
Starting point task number	0.0159	0.0331***	0.043	0.0184	0.0216*
	(1.70)	(4.27)	(1.70)	(1.57)	(2.15)
Risk neutral task number	-0.0577*	-0.379***	-0.0539*	-0.0862**	-0.170***
	(-2.12)	(-4.08)	(-1.99)	(-3.03)	(-4.74)
Risk neutral task number, squared		0.0146			
		(1.17)			
Starting point taskno-Risk neutral taskno			-0.00296		
			(-0.37)		
Dummy for starting point in risk loving zone				1.099***	0.963***
				(4.17)	(6.16)
Enumerator FE	No	No	No	No	Yes
Constant	1.202***	1.870***	1.174***	1.293***	1.777***
	(8.70)	(13.16)	(8.49)	(10.79)	(9.93)
Prelec 2 alpha					
Page number	0.0164***	0.0140***	0.0161***	0.0165***	0.0156***
	(3.52)	(3.67)	(3.49)	(3.72)	(4.38)
Starting point task number	0.00347*	0.00139	0.0120*	0.00487**	0.00472**
	(2.22)	(0.98)	(2.41)	(2.79)	(2.95)
Risk neutral task number	0.0796***	0.289***	0.0809***	0.0803***	0.0798***
	(19.41)	(15.28)	(19.53)	(19.74)	(20.33)
Risk neutral task number, squared		-0.0334***			
		(-10.13)			
Starting point taskno-Risk neutral taskno			-0.00437*		
			(-2.06)		
Dummy for starting point in risk loving zone				-0.552***	-0.483***

	(1)	(2)	(3)	(4)	(5)
				(-8.85)	(-6.92)
Enumerator FE	No	No	No	No	Yes
Constant	0.337***	0.0549	0.336***	0.331***	0.228***
	(14.03)	(1.78)	(13.97)	(14.63)	(8.75)
Prelec 2 beta					
Page number	0.0427**	0.0428***	0.0410**	0.0421**	0.0401***
	(2.79)	(3.30)	(2.73)	(3.02)	(3.79)
Starting point task number	-0.0149**	-0.0229***	-0.0368***	-0.0161*	-0.0160**
	(-3.17)	(-5.54)	(-3.65)	(-2.52)	(-3.19)
Risk neutral task number	-0.00363	0.230***	-0.00567	0.00692	0.0376**
	(-0.38)	(6.09)	(-0.60)	(0.70)	(3.06)
Risk neutral task number, squared		-0.0222***			
		(-5.25)			
Starting point taskno-Risk neutral taskno			0.00238		
			(1.17)		
Dummy for starting point in risk loving zone				-0.276**	-0.248***
				(-2.63)	(-5.15)
Enumerator FE	No	No	No	No	Yes
Constant	0.772***	0.362***	0.789***	0.733***	0.529***
	(10.77)	(6.51)	(10.94)	(11.92)	(6.11)
DFT Noise					
Page number	0.00429	0.00108	0.00398	0.00464	0.00147
	(1.11)	(0.35)	(1.04)	(1.22)	(0.50)
Starting point task number	0.0112***	0.00962***	0.0187***	0.0127***	0.0133***
	(6.38)	(5.43)	(6.01)	(6.26)	(7.35)
Risk neutral task number	-0.009***	-0.156***	-0.009***	-0.008***	-0.0004
	(-4.32)	(-13.92)	(-4.28)	(-3.85)	(-0.15)
Risk neutral task number, squared		0.0175***			
		(11.69)			
Starting point taskno-Risk neutral taskno			-0.00145*		
			(-2.15)		
Dummy for starting point in risk loving zone				0.158*	0.104
				(1.98)	(1.84)
Enumerator FE	No	No	No	No	Yes

Constant	0.414***	0.665***	0.415***	0.411***	0.326***
	(23.54)	(28.94)	(23.75)	(24.46)	(11.60)
Number of observations	116577	116577	116577	116577	116577

Note: t statistics in parentheses, * p<0.05, ** p<0.01, *** p<0.001

Table 11A. Correlations between number of durable assets and predicted risk preference parameters from models with 10 CL vs. 12 CL with DFT and CU specification. Poisson panel models with enumerator random effects, without and with predicted noise parameters.

durassetno	(1) DFT10	(2) DFT10+e	(3) DFT12	(4) DFT12+e	durassetno	(5) CU10	(6) CU10+e	(7) CU12	(8) CU12+e
Rrdu-DFT10	-2.038***	-3.406***			Rrdu-CU10	-2.560***	-3.447***		
	(0.296)	(0.285)				(0.338)	(0.326)		
Ardu-DFT10	-0.392	-5.949***			Ardu-CU10	-0.547	-4.247***		
	(0.481)	(0.611)				(0.408)	(0.387)		
Brdu-DFT10	-0.646	-4.754***			Brdu-CU10	-2.390***	-5.962***		
	(0.684)	(0.529)				(0.568)	(0.488)		
Nrdu-DFT10		39.30***			Nrdu-CU10		116.4***		
		(3.139)					(9.253)		
Rrdu-DFT12			-8.534***	-8.257***	Rrdu-CU12			-5.892***	-5.673***
			(0.497)	(0.449)				(0.411)	(0.332)
Ardu-DFT12			-6.404***	-10.13***	Ardu-CU12			-5.917***	-8.746***
			(0.450)	(0.701)				(0.466)	(0.666)
Brdu-DFT12			-11.57***	-11.36***	Brdu-CU12			-10.29***	-10.27***
			(0.740)	(0.624)				(0.740)	(0.562)
Nrdu-DFT12				21.79***	Nrdu-CU12				80.52***
				(2.497)					(8.650)
Constant	1.503	-3.959**	21.67***	16.20***		3.463***	-3.995**	17.24***	9.706***
	(0.960)	(1.283)	(1.291)	(1.368)		(0.781)	(1.337)	(1.155)	(1.434)
Lnalpha	-2.026	1.401***	-1.932	0.730		-2.052*	1.482***	-1.896	1.050*
	(1.069)	(0.319)	(1.019)	(0.519)		(1.045)	(0.345)	(0.992)	(0.443)
N	945	945	945	945	N	945	945	945	945

Note: Dependent variable: durassetno=Number of durable assets. Bootstrapped standard errors in parentheses, resampling enumerators, 500 replications. * p<0.05, ** p<0.01, *** p<0.001. Rrdu=CPRRA parameter, Ardu=Prelec 2 alpha parameter, Brdu=Prelec 2 beta parameter, Nrdu=Noise parameter. All models with experimental enumerator fixed effects.

Table 12A. Correlations between livestock endowment and predicted risk preference parameters from models with 10 CL vs. 12 CL with DFT and CU specification. Panel tobit models with enumerator random effects, without and with predicted noise parameters.

TLU	(1) DFT10	(2) DFT10+e	(3) DFT12	(4) DFT12+e	TLU	(5) CU10	(6) CU10+e	(7) CU12	(8) CU12+e
Rrdu-DFT10	-4.132***	-3.982			Rrdu-CU10	-5.772***	-5.658**		
	(0.664)	(2.256)				(0.958)	(2.050)		
Ardu-DFT10	0.588	1.052			Ardu-CU10	0.799	1.014		
	(1.548)	(6.435)				(1.398)	(4.767)		
Brdu-DFT10	-1.450	-1.129			Brdu-CU10	-5.043***	-4.759		
	(1.367)	(5.369)				(1.356)	(5.581)		
Nrdu-DFT10		-4.029			Nrdu-CU10		-9.699		
		(46.130)					(142.000)		
Rrdu-DFT12			-21.98***	-21.96***	Rrdu-CU12			-14.68***	-14.69***
			(2.342)	(2.695)				(1.637)	(1.874)
Ardu-DFT12			-15.89***	-15.30*	Ardu-CU12			-12.37***	-12.02*
			(1.580)	(6.135)				(1.616)	(5.042)
Brdu-DFT12			-29.73***	-29.71***	Brdu-CU12			-25.12***	-25.11***
			(3.345)	(4.041)				(2.935)	(3.758)
Nrdu-DFT12				-3.594	Nrdu-CU12				-10.480
				(28.880)					(102.700)
Constant	2.293	3.286	55.32***	56.33***		6.156***	7.076	41.05***	42.24***
	(1.987)	(8.378)	(5.759)	(6.573)		(1.775)	(10.510)	(4.533)	(8.424)
Sigma_u	0.474	0.000	0.455	0.000		0.426	0.000	0.471	0.000
	(0.272)	(2.850)	(0.246)	(2.465)		(0.251)	(3.415)	(0.261)	(2.796)
Sigma_e	2.523***	2.535***	2.279***	2.293***		2.484***	2.497***	2.308***	2.322***
	(0.157)	(0.199)	(0.163)	(0.192)		(0.156)	(0.235)	(0.164)	(0.207)
N	945	945	945	945		945	945	945	945

Note: Dependent variable: TLU=Livestock endowment (Tropical livestock units). Bootstrapped standard errors in parentheses, resampling enumerators, 500 replications. * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Rrdu=CPRRA parameter, Ardu=Prelec 2 alpha parameter, Brdu=Prelec 2 beta parameter, Nrdu=Noise param.

Appendix 2

Risk preference Experiments

Instructions to experimental enumerators:

Explanation for risk experiments with money:

Risk of starting point bias: Do as with the time preference series:

a. Randomize the task you start with within each series (throw the die).
b. This should be part of the pre-making of the questionnaires before you start.
c. Next, you move towards the end point in the direction you expect a switch to check whether you get it.
d. Narrow in quickly on the switch point by going to the middle task between the last two tasks that was assessed and within which the switch point is located (if consistent preferences are observed).

With real money and varying probabilities of low and high outcomes

a. The probability of low (bad) outcome for the risky prospect varies from game to game.
b. You will identify the certain outcome that makes the respondent switch (switch point) between preferring the risky prospect to preferring the certain outcome.

Instructions to respondents:

1. *You have the choice between a risky prospect which has a high (good) or a low (bad) outcome.*
2. *There is in each game a certain probability (chance) of low (bad) outcome for the risky prospect.*
3. *One of the gambles gives you ETB 100 if you are lucky and ETB 0 if you are unlucky and a chance/probability, say one out of ten (10% chance) of low (bad) outcome.*
4. *You have the choice between this and a certain amount.*
5. *We vary the certain amount till you switch from preferring one or the other.*
6. *To determine whether you are lucky we will use a 20-sided die.*

Random sampling for payout:

a. One of the experiments below will be randomly sampled for real payout.
b. Your choice in the randomly sampled task in that game will be your payout.
c. Your choices will there affect the outcome.
d. Therefore, think carefully about your preferred choices.

The order of the 5 pages with risk preference experiments was randomized to reduce bias

Example of MPL for risk preference elicitation:

				Risk preference series 6				
S. no.	Start point	Task no.	Prob-bability of bad outcome	Low outcome, **ETB**	High outcome, **ETB**	Choice	Certain amount, **ETB**	Choice
6		1	2/10	20	100		100	
6		2	2/10	20	100		95	
6		3	2/10	20	100		90	
6		4	2/10	20	100		85	
6		5	2/10	20	100		80	
6		6	2/10	20	100		75	
6		7	2/10	20	100		70	
6		8	2/10	20	100		65	
6		9	2/10	20	100		60	
6		10	2/10	20	100		50	

Payment for Risk preference games:

a. Use 20-sided die (in cup with cartoon) to identify which of the 10 risk series that will be used for payout (die numbers 1 and 11 for risk series 1, die numbers 2 and 12 for risk series 2, etc.), and similarly for the choice of Task (row) in the risk series identified.

b. This should be done for each at the end of all games, while they are sitting at their place. Nobody should move from their spot till all have completed (no spectators allowed for each).

c. Ensure privacy during the whole process.

d. If some complete before others, they should not come close to or interact with those who have not yet completed.

e. You use the Prospect they have chosen for that task, the risky prospect or the certain amount depending on their choice in that specific task.

f. If they have chosen the risky prospect you identify the probability of Good (High) and Bad (Low) outcomes and assign die numbers to each, e.g. 30% probability of Good outcome in Risk series 3 game implies that you assign die numbers 1, 2, 3, 11, 12, and 13 to the low payout and the remaining die numbers to high (good) payout.

g. The die has to be shaken under the cup only once to determine the number and identify whether they lost or won.

h. If they for the randomly identified task chose the certain amount, you give them that certain amount.

Payment in risk preference experiments:

1. Risk series randomly assigned for payout:_____
2. Task row randomly assigned for payout:_____
3. Identify whether the Respondent had chosen the Risky Prospect (=1) or the Certain Amount (=2) for that Task: Prospect chosen (circle):___ 1_____2_
4. If the certain amount was chosen, write the amount below as amount recived.
   Amount recieved: _____

5. If the risky prospect was chosen, assign die numbers to low and high payouts based on the probabilities in the identified risk series.

   Die number outcome:_____

   Implication (circle): 1=High outcome, 0=Low outcome

   Amount:_____

Signature for amount received:_____

Loss Aversion Choice List: It was real and played after the other risk CLs

LA1				Prospect A: Outcomes in ETB				Prospect B: Outcomes in ETB			
Choice series no	Start point	Task no.	Prob. of bad outcome %	Win	Loss		Choice	Win	Loss		Choice
11		1	**50**	50	-10			60	-40		
11		2	**50**	30	-10			60	-40		
11		3	**50**	20	-10			60	-40		
11		4	**50**	10	-10			60	-40		
11		5	**50**	5	-10			60	-40		
11		6	**50**	5	-10			60	-30		
11		7	**50**	5	-15			60	-30		
11		8	**50**	5	-15			60	-25		
11		9	**50**	5	-15			60	-20		

Mark the task that was sampled to be real: Die outcome=**Task no**:_____

Die outcome: Game outcome (circle): 1=Win: die 11-20, 0=Loss: die 1**-10**

Outcome of the game: Amount lost:_____ Amount won:_____

Signature of player:_____

Table A14. Correlations between loss aversion, risk preferences, noise and other variables

	Poisson	XtPoisson, RE	XtPoisson, FE	XtPoisson, FE	XtPoisson, FE
	(1)	(2)	(3)	(4)	(5)
LA stp	0.0166**	0.0166**	0.0196***	0.025	0.025
	(0.005)	(0.005)	(0.005)	(0.029)	(0.029)
LA stp, squared				0.000	-0.001
				(0.003)	(0.003)
Rrdu-CU12	-0.023	-0.023	-2.921	-11.410	-11.360
	(0.135)	(0.134)	(3.941)	(6.082)	(6.054)
Ardu-CU12	-0.199	-0.199	7.359	31.350	31.220
	(0.347)	(0.345)	(10.790)	(17.120)	(17.060)
Brdu-CU12	-0.018	-0.018	-5.818	-24.010	-23.910
	(0.223)	(0.222)	(8.427)	(13.040)	(12.980)
Nrdu-CU12	-0.659	-0.659	24.550	-43.840	-42.510
	(3.297)	(3.282)	(36.000)	(64.440)	(64.440)
Male, dummy	-0.0733**	-0.0733**	-0.194	-0.663*	-0.662*
	(0.026)	(0.025)	(0.210)	(0.333)	(0.332)
Age, years				0.004	0.004
				(0.004)	(0.004)
Education, years completed				-0.174	-0.172
				(0.114)	(0.114)
Tropical livestock units					0.010
					(0.011)
Durable assets, number					-0.010
					(0.016)
Enumerator FE	Yes	Yes	Yes	Yes	Yes
Constant	1.962***	1.962***			
	(0.355)	(0.353)			
ln alpha		-17.690			
		(12.760)			
AIC	3614.9	3616.9	2910.7	2915.0	2918.5
BIC	3700.1	3706.9	2991.1	3009.6	3022.6
N	843	843	841	841	841

Note: Dependent variable: Loss aversion rank. Model (1) is a Poisson model with robust standard errors, model (2) is a panel Poisson model with youth group random effects and cluster robust standard errors, models (3)-(5) are panel Poisson models with youth group fixed effects and cluster robust standard errors. * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Rrdu=CPRRA parameter, Ardu=Prelec 2 alpha parameter, Brdu=Prelec 2 beta parameter, Nrdu=Noise parameter, LA-stp is random starting point row in loss aversion CL.