COPING WITH RISK IN POOR RURAL ECONOMIES

A thesis submitted for the degree of
D.PHIL IN ECONOMICS

by

GAUTAM KALANI
MSc BA

HERTFORD COLLEGE,
UNIVERSITY OF OXFORD.

HILARY 2013

Approximate length:
95,000 words

Supervisor:
Professor Marcel Fafchamps

For my parents

Coping with Risk in Poor Rural Economies
Gautam Kalani, Hertford College
D.Phil in Economics
Hilary 2013

ABSTRACT

Rural inhabitants of developing countries face extraordinarily risky environments, and decision-making under risk has crucial implications for the welfare of the rural poor. Therefore, obtaining a better understanding of the behaviour under risk of low-income populations is a vital step in the comprehension of human behaviour, and is important for effective policy design and evaluation, as well as for shedding light on production, investment and technology adoption decisions.

In Chapter One, I analyze data collected from a laboratory experiment involving poor subjects in rural Ethiopia, in order to determine which decision models (and corresponding risk preferences) best describe the decision-making under risk of inhabitants. I find that expected utility theory (EUT) does not provide a good overall description of the decisions made by participants in the experiment; instead, there is evidence of probability weighting and loss aversion, implying that rank-dependent and reference-dependent choice models are more likely to represent the true latent decision-making process of subjects.

In Chapter Two, I analyze combined experimental and survey data from rural Ethiopia in order to evaluate the determinants of risk preferences as well as assess the degree of asset integration in experimental decisions. Analyzing both EUT and non-EUT decision models and using an instrumental variable strategy, I find that household wealth negatively affects both risk aversion and loss aversion, but independent background risk has no effect on risk preferences. Further, I find evidence of narrow framing, as opposed to asset integration, suggesting that participants make decisions in the experiment in isolation from outside wealth.

In Chapter Three, I analyze experimental data from Brazil to evaluate whether subjects understand decision problems that use the complex Multiple Price List (MPL) elicitation procedure, and to determine which decision models best describe observed choices. I find that the MPL decision problems of the experiment enable a finer characterization of risk preferences as compared to Ordered Lottery Selection problems (used in the Ethiopian experiment). However, I find that a significant fraction of choice patterns in the MPL problems are intransitive, and the evidence indicates that subjects did not properly understand the decision problems and thus observed choices do not reveal true risk preferences. Therefore, the relatively complex MPL procedure may not be suitable for experiments conducted with poorly-educated subjects in developing country settings.

Chapter Four presents a theory outlining the relationship between rational demand for index insurance – for which the net transfer between insurer and policyholders depends only on a publicly verifiable index – and wealth. Further, the validity of this theory is tested using the experimental data from Ethiopia. In line with the theoretical model presented, due to basis risk and actuarially unfair premiums, demand for index insurance is hump-shaped – first increasing then decreasing – in wealth. The results indicate that the low take-up of this product observed among the poorest (and most risk averse) individuals in recent field studies may result from rational choice rather than credit constraints or poor decision-making.

ACKNOWLEDGEMENTS

This thesis would not have been possible without the generous assistance of many people. Most importantly, I must thank my supervisor, Professor Marcel Fafchamps, who has helped me greatly throughout my M.Sc and D.Phil research. His advice, guidance and feedback has been invaluable to my research and has shaped the way I think and write. It has truly been a privilege to work with him.

Many others have provided valuable feedback and suggestions. I would like to thank Daniel Clarke, Simon Quinn, Stefan Dercon, Sujoy Mukerji, Glenn Harrison, James Fenske, Michele Belot, Francis Teal, Catherine Porter, Riccardo Lucchetti, Daniel Osgood, Maria Velez, Alexander Pfaff, Rocco Macchiavello, Abigail Barr, Erlend Berg, Ruth Vargas Hill and Dan Stein.

Funding for this research has, at various stages, been provided by the Economics Department at the University of Oxford, Hertford College (Oxford) and Amherst College (USA).

Finally, I thank my fiancé, parents, family and friends for their love, support, encouragement and patience, both during this research process and for the many years before it. I will always be very grateful for their help.

STATEMENT OF AUTHORSHIP

Chapters One, Two and Three were sole authored. Chapter Four was coauthored with Daniel Clarke. The experiment was designed and run by Daniel. Broadly, the theoretical analysis was led by Daniel and the empirical analysis was led by Gautam. Specifically, the Introduction and Conclusion were joint work, Section 2 was led by Daniel and the remainder was led by Gautam.

TABLE OF CONTENTS

INTRODUCTION

Motivation. Rural inhabitants of developing countries face extraordinarily risky environments and are routinely exposed to large fluctuations in income (Fafchamps 2003). Further, formal institutions to deal with risk in these regions are virtually non-existent; while numerous informal mechanisms have been developed to deal with risk, there is substantial evidence that poor rural households – who are involved primarily in small-scale agriculture – are still very vulnerable to shocks (Townsend 1994, Barr and Genicot 2008). Meanwhile, the considerable adverse impact of uninsured risk on household welfare in rural regions has been well documented. For example, Dercon (2004), analyzing data from Ethiopia, finds that rural inhabitants exposed to severe weather shocks have significantly lower long-term income growth; further, even temporary shocks can reduce investment in human capital or leave households vulnerable to poverty traps through the sale (or loss) of productive assets and depletion of buffer stocks (Mosley and Verschoor 2005, Hill et al. 2011). Since most poor rural inhabitants live close to the subsistence level, the threat and impact of these shocks are magnified – indeed, risk is more prevalent, a graver threat and a greater concern in these regions than in developed economies (Besley 1995, Fafchamps 2003).

Given the significant exposure to risk, decision-making under risk and risk preferences determine important economic choices and have crucial implications for the welfare of the rural poor. Additionally, appropriately classifying behaviour under risk is crucial for effective policy design, implementation and evaluation (Harrison et al. 2010). Thus, obtaining a better understanding and characterization of the decision-making under risk of low-income populations, which is the focus of this thesis, is a vital step in the comprehension of human behaviour.

Outline and contribution. This thesis comprises four chapters, each of which is presented as a distinct research paper. It investigates the behaviour under risk of poor rural inhabitants using combined data from laboratory experiments and surveys conducted in Ethiopia and Brazil. Broadly speaking, it addresses three important topics – understanding the decision-making under risk (and the determinants of risk preferences) of individuals in developing countries, understanding decisions regarding the take-up of new financial products such as index insurance[1], and investigating the methodological issue of suitable risk elicitation procedures for developing countries. Thus, it contributes to the literature on development economics, experimental economics, and the intersection of the two, labelled experimental development economics.

In Chapter One, I analyze data from an experiment involving poor subjects in rural Ethiopia, in order to determine which decision models (and corresponding risk preferences) best describe the decision-making under risk of inhabitants. The experiment, conducted in 2009, involved 378 subjects from seven sites of the Ethiopian Rural Household Survey (ERHS). It included a benchmark decision problem, framed in the abstract, and four framed insurance decision problems. The decision problems used the Ordered Lottery Selection procedure popularized by Binswanger (1980, 1981) – in which subjects choose one out of multiple options (in this case, six) – to elicit risk preferences.

While most studies on the subject assume a particular decision-making process *a priori*, or analyze just one or two competing theories of choice, this chapter analyzes a wide range of models describing decision-making under risk, including expected utility theory (EUT), rank-dependent utility (RDU) and cumulative prospect

[1] In an index insurance arrangement, the net transfer between insurer and policyholders depends only on a publicly verifiable index. For example, in the case of rainfall (or weather) indexed insurance, policyholders receive an insurance payout if the rainfall measured at a local weather station falls below a certain level.

theory (CPT). These models are estimated using subjects' choices in the Ethiopian experiment, thus allowing the data to determine which theory (or theories) provides the best description of decision-making under risk. While EUT is still considered by many to be the dominant theory of decision-making under risk in economics, numerous experimental studies both in developed and developing countries have found strong evidence of loss aversion and probability weighting, with decisions conforming better to non-EUT models such as RDU and CPT (Hey and Orme 1994, Tanaka et al. 2010). As a result, Humphrey and Verschoor (2004) note that future research on the subject should focus not only on expected utility theory and the magnitude of the risk aversion coefficient, but also on investigating the structure of preferences and the decision-making process of individuals. This chapter therefore builds on the work of Humphrey and Verschoor (2004) and Harrison et al. (2010), who test the validity of different decision models using experimental data from India, Ethiopia and Uganda.

In order to analyze the different decision theories, I assume that choices are generated by a model of "stochastic choice with deterministic preferences" – that is, a deterministic preference-maximizing model with an added error structure – and use a structural model combined with a maximum likelihood estimation technique (following the estimation strategy pioneered by Camerer and Ho 1994). I find that expected utility theory does not provide a good overall description of the decisions made by participants in the experiment. Instead, for the insurance decision problems, I find evidence of probability weighting (specifically, "S-shaped" probability weighting) and loss aversion, implying that rank-dependent and reference-dependent choice models (such as RDU and CPT) are more likely to represent the true latent decision-making process of participants. Further, probability weighting can also explain the level of insurance take-up in the index insurance decision problems of the experiment; this level is higher than would be optimal for any expected utility maximizer with preferences satisfying risk aversion and constant relative risk aversion (and in fact, for any expected utility maximizer with preferences satisfying risk aversion and decreasing absolute risk aversion). Research on the demand for index insurance, an innovative product with large potential welfare benefits, in developing countries is still in its infancy, and, in particular, this topic has received relatively little attention in the experimental literature.

In Chapter Two, I extend these results by evaluating the determinants of risk preferences of individuals in rural Ethiopia and assessing the degree of asset integration in experimental decisions.[2] In order to do this, I combine data from the Ethiopian experiment with survey data from the ERHS – this combined dataset provides fertile ground to investigate the influence of an extensive set of demographic and socioeconomic variables on decision-making and risk-taking in the experimental decision problems. In particular, I focus on jointly establishing the impact of two economic factors – wealth and background risk – on risk attitudes.

In order to obtain a comprehensive understanding of behaviour under risk, it is crucial to evaluate the determinants of risk attitudes, in addition to estimating and characterizing the risk preferences themselves. While a growing literature argues that risk preferences are related to household economic circumstances, the existing empirical evidence on the direction and strength of the relationship between wealth and risk attitudes is far from conclusive (Liu 2008). Additionally, most studies that analyze risk-taking focus on individuals' behaviour when faced with a single risky decision. However, in reality, individuals are simultaneously exposed to a variety of risks rather than a single risk, and decisions in relation to a particular risk can often be influenced by other independent "background" risks (Guiso and Paiella 2008). Lusk and Coble (2008) note that ignoring background risk could yield misleading inferences when analyzing risk-taking behaviour.

Economic circumstances, such as wealth and background risk, are likely to be

[2] The phenomenon of integration of assets and income from all sources in the decision-making of individuals is referred to as "asset integration", while the consideration of only the prospective gains and losses associated with the current decision when making choices has been labelled "narrow framing" (Heinemann 2008).

affected by risk preferences and thus may be endogenous due to simultaneity (reverse causality) (Guiso and Paiella 2008). In order to deal with the potential endogeneity associated with these variables and obtain consistent estimates of the causal impact of wealth and background risk on risk preferences, I employ an instrumental variable (IV) strategy, using mainly village-level variables interacted with the gender of the respondent as instrumental variables. To estimate the IV models, I use two-step and Limited Information Maximum Likelihood interval estimators (described by Bettin and Lucchetti 2010), as well as a two-step version of the structural maximum likelihood estimator detailed in the work of Harrison and Rutström (2008). To my knowledge, this analysis represents the first application of these estimators to the evaluation of the determinants of risk preferences. Further, there are very few studies on the determinants of risk attitudes that jointly consider both wealth and background risk, and even fewer that suitably account for the endogeneity associated with both these variables. Indeed, this also represents the first analysis of the determinants of risk preferences in a low-income setting which involves instrumenting for both wealth and background risk.

Additionally, while most studies on the subject use single-parameter EUT models, this chapter uses both EUT and non-EUT frameworks to assess the determinants of risk aversion as well as loss aversion, building on the work of Tanaka et al. (2010). They note that research on this topic using non-EUT frameworks is crucial, given the strong existing empirical evidence that loss aversion is an important characterization of the behaviour under risk of the rural poor, and that non-EUT decision models with multiple preference parameters better fit both experimental and field data than EUT models.

I find some evidence that household wealth negatively affects both risk aversion and loss aversion. This implies that richer households are less averse to both risk and losses than poorer households, in line with the results of recent experimental studies conducted in developing countries (for example, Yesuf and Bluffstone 2009, Tanaka et al. 2010, Tanaka and Munro 2012). However, the results also indicate that independent background risk has no significant effect on the risk preferences of individuals. Thus, in line with the work of Alessie et al. (2002) and Lusk and Coble (2008), I do not find support for Gollier and Pratt's (1996) risk vulnerability hypothesis, according to which higher background risk should be associated with greater (indirect) risk aversion.

Further, I find evidence of narrow framing, as opposed to asset integration, suggesting that participants isolate decisions in the experimental lotteries from other decisions in their lives, and make choices by considering only the prospective gains and losses associated with the current decision, independent of outside wealth.[3] This analysis of asset integration using experimental data is related to the work of Harrison et al. (2007) and Heinemann (2008), and allows the data to determine the argument (or input) of the utility function.[4] Correctly identifying the argument of the utility function is crucial for the accurate estimation of risk attitudes. However, the data analyzed by Harrison et al. (2007) and Heinemann (2008) does not contain information on subjects' outside wealth – as a result, they can only test the hypothesis of narrow framing, and cannot estimate the extent of asset integration. Using data on household wealth from the ERHS, I build on their work and estimate more flexible asset integration specifications, thus obtaining robust and accurate estimates of the degree of asset integration, in addition to testing the narrow framing hypothesis. Further, most studies analyzing asset integration use data from developed country experiments, and the phenomenon of asset integration in the context of experiments conducted in developing countries has been largely ignored in the empirical literature.

Thus, the robust analysis of the extent of asset integration, joint estimation of the impact of wealth and background risk on risk preferences using suitable IVs and novel estimators, and the analysis of the determinants of loss aversion (in addition to risk aversion) represent the major empirical contributions of this chapter.

[3] The non-integration of outside wealth and the isolation of experimental decisions highlights the issue of field relevance (or external validity) of laboratory results. Concerns associated with the external validity of lab experiments have been raised by Hill and Nobles (2011) and Lybbert et al. (2010).

[4] For a utility function $U(x)$, x is referred to as the argument – it is the factor or variable (for example, income) over which utility is defined.

In Chapter Three, I analyze data from a lab experiment in Brazil which featured decision problems following the relatively complex (but widely used) Multiple Price List (MPL) elicitation procedure, in order to determine how well subjects understand this procedure, and thus assess the suitability of this method for experiments conducted in developing countries. While the Ethiopian experiment analyzed in Chapters One and Two consisted of Ordered Lottery Selection problems (in which participants make only a single choice), in each decision problem of the Brazilian experiment, participants were required to answer 20 binary questions indicating whether or not they would prefer to buy a particular insurance contract at different prices.

Since a MPL decision problem involves a relatively large number of choices, it enables the classification of subjects into a greater number of risk preference categories – and thus facilitates a more precise calibration of risk attitudes – than an Ordered Lottery Selection decision problem (Dave et al. 2010). This greater precision, combined with the larger number of total choices in the dataset, associated with the MPL procedure makes it more suitable for obtaining accurate point estimates of preference parameters using econometric methods (Andreoni and Sprenger 2012, Charness et al. 2013). As a result, in Chapter One, using data from the five Ordered Lottery Selection problems of the Ethiopian experiment, it is not possible to identify and accurately estimate the parameters of certain decision models.

In Chapter Three, I explore these identification and estimation issues, in addition to the question of comprehension of MPL decision problems; using the MPL data from the Brazilian experiment (which involved 266 participants), I estimate all the major decision models analyzed in Chapter One (including EUT, RDU and CPT). Further, this analysis facilitates the determination of which decision models (and corresponding preference parameters) best describe the behaviour under risk of individuals in the Northeast region of Brazil, which is characterized by drought, low agricultural productivity and persistent poverty (Broad et al. 2007).

I do find that the MPL procedure of the Brazilian experiment provides more power and enables the estimation of a greater range of decision models (especially non-EUT models) as compared to the Ordered Lottery Selection problems of the Ethiopian experiment.[5] However, decision problems using the MPL elicitation technique are cognitively more challenging since they involve more choices (thus requiring more complex mental calculations and placing a greater "cognitive load" on subjects) (Dave et al. 2010). Indeed, Crosetto and Filippin (2013) and Charness et al. (2013) note that the two methods differ substantially in their complexity – the MPL procedure is significantly more difficult to comprehend than the simpler Ordered Lottery Selection method, particularly for individuals with low levels of education and poor mathematical skills.

As a result, I obtain estimates of the noise parameter in Luce and Fechner error models that imply a considerable degree of randomness in choice. Additionally, I find that a substantial fraction (33%) of choice patterns in the MPL decision problems are intransitive – due to multiple or reverse switching – and thus are not consistent with any of the standard decision models, and indeed with any well-defined preferences over risk. Barr (2007) notes that the presence of such inconsistent choice patterns could indicate a lack of understanding of the decision problems, random errors in decision-making, or genuine indifference between the alternatives. An important contribution of this chapter is the determination of which of these factors is the most likely cause of the inconsistent patterns observed in the data – this issue has not been analyzed in depth in the experimental literature.

Using data on participant characteristics collected in conjunction with the experiment to estimate Probit and Ordered Probit specifications, I find that participants who did not attend college are significantly more likely to have intransitive choice

[5] The estimates of most preference parameters (for example, the probability weighting parameter) differ significantly when using data from the Brazilian and Ethiopian experiments. This implies quantitatively and qualitatively different behaviour under risk for the two samples of subjects, highlighting the difficulty in extending experimental results from one setting to another.

patterns, indicating that the observed inconsistencies could be related to low cognitive ability. Additionally, using the constant error rate analysis (developed by Harless and Camerer 1994), I reject the hypothesis that the inconsistent choice patterns are caused by random ("trembling hand") errors. To my knowledge, the constant error rate analysis has not been previously conducted in this context using MPL data involving a relatively large number of choices in each decision problem, such as in the Brazilian experiment. Further, I find that indifference between alternatives in the decision problems is unlikely to cause the significant fraction of inconsistent patterns observed in the data.

Using Vuong tests and mixture model specifications (as described by Harrison and Rutström 2009), I find that an alternative decision model – which does not assume transitivity in choice and is not based on well-defined preferences over risk – better fits the data than any of the standard models considered. Thus, the results indicate that the choices of participants in the MPL decision problems may not reveal true risk preferences (which are presumably transitive), but rather provide information on subjects' preferences about how to play the "games" in a way that is not completely determined by preferences over risk. Further, it is likely that participants use the metric or heuristic implied by this alternative model for making decisions – thus essentially making choices in a psuedo-random fashion – due to confusion and poor understanding of the decision problems; this is the most likely cause of the significant fraction of intransitive choice patterns observed in the data, rather than indifference or "trembling hand" errors.[6] This highlights the issue of poor understanding associated with the MPL procedure. Thus, the results in this chapter, in line with those of Charness and Viceisza (2012), indicate that the relatively complex MPL procedure may not be suitable for eliciting consistent choices that accurately reflect true risk preferences in developing country settings characterized by low education levels; instead, simpler procedures, such as the Ordered Lottery Selection method, are likely to be preferred

[6] Humphrey and Verschoor (2004) also note that experiment participants may use simplifying heuristics in decision-making when they face difficulty in comprehending decision problems.

for experiments conducted in these settings.

This is the major methodological contribution of Chapter Three; work on this topic is part of the methodological effort to improve experimental design and enhance the measurement of risk preferences (Dave et al. 2010). Lybbert et al. (2010) note that complexity and participant understanding of decision problems are crucial factors to consider when conducting experiments in developing countries with individuals having limited formal education. However, the analysis of these issues has been given relatively little attention in the experimental development economics literature, despite the rapid increase in the number of experiments conducted in developing countries over the past decade and the dependence of task suitability on the setting and characteristics of the sample population.

This chapter therefore provides a methodological and empirical contribution, and builds on the work of Dave et al. (2010), Charness and Viceisza (2012) and Crosetto and Filippin (2013) in assessing the suitability of the MPL procedure for accurately eliciting risk preferences. However, most studies on this topic analyze data from experiments conducted in developed countries (with the exception of Charness and Viceisza 2012), and primarily consider expected utility theory. Thus, the large estimates of the noise parameter they obtain using data from MPL lotteries may be because EUT does not accurately describe the decision-making process of participants, rather than due to randomness in choice caused by poor understanding. Therefore, it is crucial to consider a wide range of different decision models in analyses of behaviour under risk, as assuming an incorrect decision-making process could yield erroneous results (Charness et al. 2013). Additionally, the above-mentioned studies do not test between the different possible causes of observed inconsistencies in choice.

Further, the results of studies (for example, Dave et al. 2010, Crosetto and Filippin 2013) that use a comparison of noise parameter estimates (in Luce and Fechner error models) to compare the extent of errors in decision-making across contexts in which the preference parameters differ could be misleading, since the estimate of the noise parameter depends also on the preference parameter values, and variation in the noise parameter does not solely capture differences in noise. Therefore, I propose a normalization procedure which enables the direct comparison of the noise parameter across contexts in which preference parameters differ – this represents another important contribution of this chapter. The results in this chapter also call into question the commonly-used assumption of transitivity in experimental choice; therefore, experiments which use variations of the MPL procedure that impose consistency in choice (for example, Liu 2008, Tanaka et al. 2010) could yield misleading inferences regarding decision-making under risk.

In order to analyze the demand for index insurance, Chapter Four – which was coauthored with Daniel Clarke – presents a theory outlining the relationship between rational demand for index insurance and wealth. This theory is then empirically tested using ERHS data combined with data on the choices of participants in one of the decision problems of the Ethiopian experiment which was framed in terms of a real-world weather indexed insurance purchase decision. We find that the demand for index insurance from rural Ethiopians is hump-shaped in wealth. That is, demand is first increasing then decreasing in wealth, with the poorest participants (who are expected to be the most risk averse) and the richest participants (who are expected to be the least averse to risk) in the experiment having the lowest demand, while the highest demand is from subjects with intermediate levels of wealth. This shape of the demand for index insurance is consistent with expected utility theory for the class of utility functions satisfying constant relative risk aversion, as outlined in the theory presented. Additionally, we do not find strong evidence that schooling, understanding of the decision problems, or quantitative literacy increase index insurance take-up. Thus, the results imply that the take-up of index insurance is "rationally" low for the poorest (most risk averse) and wealthiest (least risk averse) individuals, due to basis risk and actuarially unfair premiums respectively.[7]

Of the many risks faced by households in poor rural economies, weather risk is often cited as the most prevalent as well as the most damaging source of risk in these regions, due to its significant impact on both the level and variability of agricultural income, in addition to its short-term and long-term impact on household welfare (Rosenzweig and Binswanger 1993, Dercon 2004, Cole et al. 2009). Thus, products that provide insurance against weather shocks are expected to yield considerable welfare benefits. In the last decade, a variety of institutions have piloted the sale of weather indexed insurance – under which the net transfer between insurer and policyholders depends only on publicly observable readings from a contractual weather station – to poor farmers. Giné et al. (2008) note that index insurance has large potential risk management and welfare benefits, as well as greater transparency, lower transaction costs and lesser moral hazard and adverse selection problems than traditional crop insurance arrangements. However, voluntary purchase of these products in field studies conducted by Giné et al. (2008) and Cole et al. (2009) in rural India and by Hill et al. (2011) in Ethiopia has been relatively low, and evidence indicates that the poorest and most risk averse farmers have the lowest demand.

The results in this chapter imply that the low take-up observed, particularly among the poorest (and most risk averse) individuals, in these studies may be a result of rational choice – due to a fundamental lack of desirability of the product – rather than barriers associated with the adoption of a new technology, such as lack of understanding, credit constraints, unwillingness to experiment with new products or poor decision-making on the part of rural consumers. Further, our theory and empirical evidence of a hump-shaped relationship between index insurance take-up and wealth contradict the prediction of Giné et al.'s (2008) "neoclassical" insurance demand model that take-up should decrease monotonically with wealth, as well as the predictions of

[7] Basis risk is defined as the risk that income from index insurance will not accurately reflect the incurred loss; in other words, it is the risk that the policyholder experiences a bad individual outcome but the index takes on a value which does not trigger the insurance payout.

technology adoption models, which predict that take-up should increase monotonically with wealth and which Hill et al. (2011) believe to be the most suitable for describing the take-up of index insurance. Thus, the results indicate that the theoretical model outlined in this chapter may better inform the shape of index insurance take-up.

This chapter therefore contributes to the theoretical and empirical literature on the demand for index insurance in developing countries, and analyzes the correlates of index insurance take-up. It is related to the work of Giné et al. (2008), Cole et al. (2009) and Hill et al. (2011), who also analyze the take-up of weather indexed insurance and its correlates. Research on this topic is crucial, as it can help shed light on the drivers of index insurance demand and address the causes of low take-up observed in pilot programs (Hill et al. 2011).

However, while acknowledging that index insurance differs significantly from indemnity insurance, these studies do not consider distinct theoretical models of take-up that are specific to index insurance and account for basis risk, despite evidence indicating that basis risk plays a significant role in index insurance purchase decisions. Further, they do not test or control for the non-linear effect of wealth on the demand for index insurance, as predicted by the theoretical model in this chapter. These studies analyze field and survey data, rather than data from a lab experiment – therefore, they do not possess information on the exact joint probability distribution of the index and losses, which is crucial for drawing conclusions and making clear normative statements about index insurance demand (Clarke 2011). In addition, Hill et al.'s (2011) analysis is based on data from questions regarding a hypothetical insurance contract with hypothetical payoffs; problems relating to the use of questions involving hypothetical payoffs have been well documented, and risk preferences elicited in non-incentivized tasks may not reflect true risk attitudes (particularly in the domain of financial decision-making) (see, for example, Holt and Laury 2002, Charness et al. 2013).

Therefore, this thesis makes empirical, methodological and theoretical contributions to the literature on decision-making under risk, index insurance demand and experimental design in the context of poor rural economies.

References

Alessie, RJM, S. Hochgürtel, and A.H.O. Soest, "Household portfolios in the Netherlands," in L. Guiso, M. Haliassos, and T. Jappelli, eds., *Household Portfolios*, MIT Press, 2002, pp. 341–388.

Andreoni, James and Charles Sprenger, "Estimating Time Preferences from Convex Budgets," *The American Economic Review*, 2012, *102* (7), 3333–3356.

Barr, A. and G. Genicot, "Risk Sharing, Commitment, and Information: An Experimental Analysis," *Journal of the European Economic Association*, 2008, *6* (6), 1151–1185.

Barr, Abigail, "Attitudes to Risk in Ghana: Field Manual," 2007.

Besley, T., "Nonmarket institutions for credit and risk sharing in low-income countries," *The Journal of Economic Perspectives*, 1995, *9* (3), 115–127.

Bettin, G. and R. Lucchetti, "Interval regression models with endogenous explanatory variables," *Empirical Economics*, 2010, pp. 1–24.

Binswanger, H.P., "Attitudes toward risk: Experimental measurement in rural India," *American Journal of Agricultural Economics*, 1980, *62* (3), 395–407.

——, "Attitudes toward risk: Theoretical implications of an experiment in rural India," *Economic Journal*, 1981, *91*, 867–890.

Broad, K., A. Pfaff, R. Taddei, A. Sankarasubramanian, U. Lall, and F. de Assis de Souza Filho, "Climate, stream flow prediction and water management in northeast Brazil: societal trends and forecast value," *Climatic Change*, 2007, *84* (2), 217–239.

Camerer, C.F. and T.H. Ho, "Violations of the betweenness axiom and nonlinearity in probability," *Journal of Risk and Uncertainty*, 1994, *8* (2), 167–196.

Charness, G. and A. Viceisza, "Comprehension and risk elicitation in the field: Evidence from rural Senegal," *IFPRI Discussion Paper*, 2012.

Charness, Gary, Uri Gneezy, and Alex Imas, "Experiential methods: Eliciting risk preferences," *Journal of Economic Behavior & Organization*, 2013.

Clarke, D.J., "A Theory of Rational Demand for Index Insurance," Department of Economics Discussion Paper Series 572, University of Oxford 2011.

Cole, Shawn A., Xavier Giné, Jeremy B. Tobacman, Petia B. Topalova, Robert M. Townsend, and James I. Vickery, "Barriers to Household Risk Management: Evidence from India," Working Paper 09-116, Harvard Business School 2009.

Crosetto, Paolo and Antonio Filippin, "A Theoretical and Experimental Appraisal of Five Risk Elicitation Methods," Technical Report, Friedrich-Schiller-University Jena, Max Planck Institute of Economics 2013.

Dave, C., C.C. Eckel, C.A. Johnson, and C. Rojas, "Eliciting risk preferences: When is simple better?," *Journal of Risk and Uncertainty*, 2010, pp. 1–25.

Dercon, S., "Growth and shocks: evidence from rural Ethiopia," *Journal of Development Economics*, 2004, *74* (2), 309–329.

Fafchamps, M., *Rural Poverty, Risk and Development*, Edward Elgar Publishing, 2003.

Giné, X., R. Townsend, and J. Vickery, "Patterns of rainfall insurance participation in rural India," *The World Bank Economic Review*, 2008, *22* (3), 539.

Gollier, C. and J.W. Pratt, "Risk vulnerability and the tempering effect of background risk," *Econometrica: Journal of the Econometric Society*, 1996, pp. 1109–1123.

Guiso, L. and M. Paiella, "Risk aversion, wealth, and background risk," *Journal of the European Economic Association*, 2008, *6* (6), 1109–1150.

Harless, D.W. and C.F. Camerer, "The predictive utility of generalized expected utility theories," *Econometrica: Journal of the Econometric Society*, 1994, pp. 1251–1289.

Harrison, G.W. and E.E. Rutström, "Risk aversion in the laboratory," in J. C. Cox and G. W. Harrison, eds., *Research in Experimental Economics*, Vol. 12, Emerald Group Publishing Limited, 2008, pp. 41–196.

____ and ____ , "Expected utility theory and prospect theory: One wedding and a decent funeral," *Experimental Economics*, 2009, *12* (2), 133–158.

____ , J.A. List, and C. Towe, "Naturally occurring preferences and exogenous laboratory experiments: A case study of risk aversion," *Econometrica*, 2007, *75* (2), 433–458.

____ , S.J. Humphrey, and A. Verschoor, "Choice under Uncertainty: Evidence from Ethiopia, India and Uganda," *The Economic Journal*, 2010, *120* (543), 80–104.

Heinemann, F., "Measuring risk aversion and the wealth effect," *Research in Experimental Economics*, 2008, *12*, 293–313.

Hey, J.D. and C. Orme, "Investigating generalizations of expected utility theory using experimental data," *Econometrica: Journal of the Econometric Society*, 1994, pp. 1291–1326.

Hill, R. V. and M. Nobles, "Flexible insurance for heterogenous farmers: results from a small scale pilot in Ethiopia," *IFPRI discussion paper 01092*, 2011.

Hill, R.V., J. Hoddinott, and N. Kumar, "Adoption of weather index insurance: Learning from willingness to pay among a panel of households in rural Ethiopia," *IFPRI discussion paper 01088*, 2011.

Holt, C.A. and S.K. Laury, "Risk aversion and incentive effects," *American Economic Review*, 2002, *92* (5), 1644–1655.

Humphrey, S.J. and A. Verschoor, "Decision-making under risk among small farmers in East Uganda," *Journal of African Economies*, 2004, *13* (1), 44.

Liu, Elaine M, "Time to change what to sow: Risk preferences and technology adoption decisions of cotton farmers in China," *Review of Economics and Statistics*, 2008.

Lusk, J.L. and K.H. Coble, "Risk aversion in the presence of background risk: Evidence from an economic experiment," *Research in Experimental Economics*, 2008, *12*, 315–340.

Lybbert, Travis J, Francisco B Galarza, John McPeak, Christopher B Barrett, Stephen R Boucher, Michael R Carter, Sommarat Chantarat, Aziz Fadlaoui, and Andrew Mude, "Dynamic Field Experiments in Development Economics: Risk Valuation in Morocco, Kenya, and Peru," *Agricultural and Resource Economics Review*, 2010, *39*, 1–17.

Mosley, P. and A. Verschoor, "Risk attitudes and the 'vicious circle of poverty'," *The European Journal of Development Research*, 2005, *17* (1), 59–88.

Rosenzweig, M.R. and H.P. Binswanger, "Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments," *The Economic Journal*, 1993, *103* (416), 56–78.

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

Tanaka, Y. and A. Munro, "Regional Variation in Risk and Time Preferences: Evidence from a Large-Scale Field Experiment in Rural Uganda," *National Graduate Institute for Policy Studies Working Paper, Tokyo*, 2012.

Townsend, Robert M., "Risk and Insurance in Village India," *Econometrica*, 1994, *62* (3), 539–591.

Yesuf, M. and R.A. Bluffstone, "Poverty, risk aversion, and path dependence in low-income countries: Experimental evidence from Ethiopia," *American Journal of Agricultural Economics*, 2009, *91* (4), 1022–1037.

IDENTIFICATION OF RISK PREFERENCES FROM EXPERIMENTAL DATA

GAUTAM KALANI*

Abstract

In this chapter, I analyze data collected from a laboratory experiment involving poor subjects in rural Ethiopia, in order to determine which decision models (and corresponding risk preferences) best describe the decision-making under risk of inhabitants. I analyze a range of different behavioural models within this experimental framework, and find that expected utility theory (EUT) does not provide a good overall description of the decisions made by participants in the experiment. Instead, for the framed microinsurance decision problems of the experiment, I find evidence of probability weighting (specifically, "S-shaped" probability weighting) and loss aversion, implying that rank-dependent and reference-dependent choice models are more likely to represent the true latent decision-making process of participants. Additionally, I find that the level of index insurance take-up in the experimental decision problems is higher than would be optimal for any expected utility maximizer with preferences satisfying risk aversion and constant relative risk aversion (and in fact, for any expected utility maximizer with preferences satisfying risk aversion and decreasing absolute risk aversion), but this can be explained by probability weighting and the non-EUT decision models considered.

JEL codes: C91, D01, D03, D81, G22, O16.

*DPhil student: Department of Economics, Centre for the Study of African Economies (CSAE) and Hertford College, University of Oxford (**gautam.kalani@economics.ox.ac.uk**). This paper is part of the work towards my DPhil thesis, which is supervised by Professor Marcel Fafchamps; the work would not have been possible without his very generous assistance. Additionally, I am grateful to Daniel Clarke and Simon Quinn, who offered useful suggestions throughout the project. I also thank Stefan Dercon, Sujoy Mukerji and Glenn Harrison for providing very useful comments on aspects of the paper.

1 INTRODUCTION

Understanding decision-making under the effect of risk is a vital step in the comprehension of human behaviour – choice under risk is a fundamental activity in all spheres of life, but it is particularly important in the socio-economic sphere. Even though economic theory now provides us with a rich set of competing models of behaviour under risk – with expected utility theory as one of the front runners – debates over the validity of these theories still rage on, with no single model emerging as a decisive descriptor of decision-making under risk (Harrison and Rutström 2009). Further, empirical questions about the suitability and applicability of these theories in different domains remain unanswered.

In particular, the characterization of decision-making under risk of poor indviduals in developing countries remains unresolved, with different studies reaching markedly different conclusions. Initial analyses – including that of Binswanger (1980, 1981), who was the first to measure risk attitudes in experimental conditions in a low-income country – were rooted in expected utility theory and focused on providing estimates of coefficients of risk aversion (Lybbert et al. 2010). However, recent evidence – particularly from Africa and Asia – has suggested that expected utility theory may not appropriately describe decision-making under risk in the developing country context (Humphrey and Verschoor 2004a, Harrison et al. 2010, Tanaka et al. 2010). This is also in line with the results from a multitude of experiments conducted in the developed world (for example, Camerer and Ho 1994, Camerer 1998). As a result, Humphrey and Verschoor (2004a) note that future research on the subject should focus not only on expected utility theory and the magnitude of the risk aversion coefficient, but also on investigating the structure of preferences and the decision-making process of individuals.

In this chapter, I analyze data from a laboratory experiment conducted in rural Ethiopia – a region characterized by poverty and rain-fed subsistence agriculture (Dercon 2004) – in order to determine which decision models best describe the decision-making under risk of inhabitants, as well as understand the important features of risk preferences. I test expected utility theory (EUT) but also assess the validity of a wide range of competing decision theories, including Quiggin's (1982) rank-dependent utility (RDU) and Tversky and Kahneman's (1992) cumulative prospect theory (CPT), which are broadly considered to be the major challengers to expected utility theory. While most studies on the subject assume a particular decision-making process (generally EUT) *a priori*, or analyze just one or two competing theories of choice, this chapter analyzes a range of different behavioural models within the same experimental framework, and also separately tests the significance of loss aversion and probability weighting. It is important to consider a wide range of decision models in analyses of behaviour under risk, as assuming an incorrect decision-making process could lead to misleading results and biased estimates of preference parameters that do not accurately reflect true risk attitudes (Charness et al. 2013).[1] Thus, I aim to impose fewer restrictions, use different estimation techniques and let the data determine which theory (or theories) provides the best description of the decision-making under risk of participants in the experiment. Additionally, I aim to accurately estimate and characterize the risk preferences of subjects in the experiment, and by extension, of individuals in this region.[2]

I find that expected utility theory does not provide a good overall description of the decisions made by participants in the experiment. Instead, for the insurance decision problems, I find evidence of probability weighting (specifically, "S-shaped" probability weighting) and loss aversion, implying that rank-dependent and reference-dependent models provide a better description of choices and thus are more likely to represent the true latent decision-making process of participants. However, I also find that certain choice models cannot be estimated and certain preference parameters cannot be identified using the Ethiopian

[1] For example, if the decision-making of participants in experiments which are analyzed using only EUT models (as if often the case) is in fact better described by non-EUT models, this could yield misleading estimates of the degree of risk aversion (Humphrey and Verschoor 2004a).

[2] Throughout this thesis, I follow the nomenclature used by Harrison and Rutström (2008) and refer to the properties (or parameters) of the utility function as risk attitudes or risk preferences.

experimental data, and discuss how this is likely a result of the experimental data at hand (specifically, the elicitation procedure and associated precision and power).

This chapter therefore builds on recent work analyzing decision-making under risk using laboratory experiments conducted in developing countries, which have increased rapidly in number over the past decade – for example, Barr (2003b) (Zimbabwe), Botelho et al. (2005) (Timor l'Este) and Dercon et al. (2011) (Kenya). In particular, it is related to the work of Humphrey and Verschoor (2004a) and Harrison et al. (2010), who use experimental data from India, Ethiopia and Uganda to determine the decision-making process and risk preferences of inhabitants.

In addition to testing between different decision theories, I analyze the choices of participants in the framed microinsurance decision problems to evaluate attitudes toward, and gauge the demand for, indemnity and index insurance products[3]; this topic has received relatively little attention in the experimental literature so far. I find that the level of insurance take-up in the index insurance decision problems is higher than would be optimal for any expected utility maximizer with preferences satisfying risk aversion and constant relative risk aversion (and in fact, for any expected utility maximizer with preferences satisfying risk aversion and decreasing absolute risk aversion). However, this phenomenon can be explained by probability weighting and the non-EUT decision models considered. Further, the results highlight the importance of lab experiments for making normative statements about the level of index insurance demand – in order to do so, precise knowledge of the joint probability distribution of losses and index claim payments is required, which is almost impossible to obtain in the field (Clarke 2011).

Most inhabitants of poor rural regions are involved in small-scale agriculture, which is subject to a great deal of risk from pests and fluctuations in weather conditions, thus causing considerable instability in income streams. Additionally,

[3] In an index insurance arrangement, the net transfer between insurer and policyholders depends only on a publicly verifiable index.

formal institutions to deal with risk in these regions are scarce, while the informal mechanisms that individuals rely on provide only limited protection against shocks (Townsend 1994, Fafchamps 2003, Barr and Genicot 2008). Further, Collier and Gunning (1999) find that farmers in rural Africa face more volatile environments than those in other parts of the world. Meanwhile, the considerable adverse impact of uninsured risk on household welfare in rural regions has been well documented. For example, Dercon (2004), analyzing data from Ethiopia, finds that rural inhabitants exposed to severe weather shocks have significantly lower long-term income growth; further, even temporary shocks can reduce investment in human capital or leave households vulnerable to poverty traps through the sale (or loss) of productive assets and depletion of buffer stocks (Mosley and Verschoor 2005, Hill et al. 2011).

Given the substantial exposure to risk, risk attitudes play a crucial role in technology adoption, investment, production, risk pooling and insurance purchase decisions, and thus impact the overall well-being of individuals in low-income regions through determining wealth accumulation and consumption (as well as income) variability (Rosenzweig and Binswanger 1993, Liu 2008, Attanasio et al. 2012). Additionally, Humphrey and Verschoor (2004a) observe that risk preferences influence the nature of numerous important interventions, such as microinsurance and microfinance programs. These interventions have the potential to significantly increase the welfare of poor rural inhabitants by raising the level of income and by reducing income risk (Fafchamps 2003). Therefore, as noted by Harrison et al. (2010), it remains a high priority to accurately estimate and characterize the risk preferences of these populations, both because decision-making under risk has considerable implications for the welfare of the rural poor, and because it is important for the effective design and evaluation of risk-mitigating and welfare-enhancing interventions.

In spite of the dense theoretical literature on decision-making under risk, it is very difficult, if not impossible, to directly test the validity of the different decision theories in the field. Further, approaches which attempt to impute risk preferences from standard household survey questions suffer from serious drawbacks (Liu 2008). On the other hand, the laboratory environment, and associated experimental methods, provides an appropriate setting for the elicitation of risk preferences; it offers a much greater degree of control, thus allowing us to observe decision-making under risk in isolation from other factors that could influence behaviour (Harrison and List 2004, Liu 2008). Indeed, Lybbert et al. (2010) note that experiments can generate insights into behaviour under risk that are impossible to obtain using standard research methods. However, Harrison et al. (2007) note that the imposition of laboratory controls may sometimes make it harder to make reliable inferences about field behaviour, which is ultimately what we are interested in.

While the use of a laboratory experiment entails a tradeoff between control and realism, the experiment analyzed in this study was designed to maximize external validity with decision problems framed as agricultural insurance purchase problems, an experimental design that yielded clear theoretical predictions and payoffs of up to one week's income. Moreover, subjects were chosen from poor households in the Ethiopian Rural Household Survey (ERHS), and some of these households would be offered real agricultural insurance policies in the subsequent two years as part of a pilot project conducted by the International Food Policy Research Institute (IFPRI) and the University of Oxford. Therefore, the participants were individuals who routinely deal with risk and make decisions about insurance purchase, and the stimuli and framing used in the experiment closely matched what they experience in the field. Harrison et al. (2007) conclude that such lab experiments conducted in the field with non-standard subject pools of developing country inhabitants represent a useful middle ground between the lab and the field, offering a significant degree of control while increasing the external validity of experimental results. Additionally, numerous experimental studies have found that risk attitudes exhibited in experimental decision problems are related to crucial economic decisions in real life (such as technology adoption and insurance purchase decisions), highlighting the importance of eliciting risk preferences in experimental settings (for example, Barr and Packard 2005, Liu 2008, Attanasio et al. 2012).

The rest of this chapter is organized as follows. The next section describes some influential models of decision-making under risk. Section 3 outlines the empirical investigation strategy involving these decision theories, while Section 4 details the experimental design and procedure. The empirical analysis is presented in Section 5, while Sections 6 and 7 investigate two interesting patterns observed in the data in the context of the empirical findings. Section 8 concludes.

2 MODELS OF DECISION-MAKING UNDER RISK

2.1 Expected utility theory

Since its inception by von Neumann and Morgenstern (1944), expected utility theory has been the dominant theory of decision-making under risk in economics (Humphrey and Verschoor 2004a). A common starting point for analyses involving experimental data is an EUT specification with a constant relative risk aversion (CRRA) utility function of the form:

$$U(x) = \begin{cases} \dfrac{x^{1-r}}{1-r} & \text{if } r \neq 1 \\ \ln(x) & \text{if } r = 1 \end{cases} \tag{1}$$

where x is the lottery prize and r is the coefficient of relative risk aversion: $r < 0$ corresponds to risk loving behaviour, $r = 0$ to risk neutral and $r > 0$ to risk averse.[4,5] Under EUT, the decision-maker weights each possible outcome $k_i \in \{1, \ldots, K\}$ in lottery i using its corresponding probability p_{k_i}, and so expected utility from lottery i (EU_i) is the sum of the probability weighted utility of each outcome in the lottery:

$$EU_i = \sum_{k_i=1}^{K} p_{k_i} U_{k_i} \tag{2}$$

The CRRA functional form, however, is quite restrictive, as it assumes that relative risk aversion is constant over the prize domain. To allow for the possibility that relative risk aversion is not constant, Saha (1993) proposes a more flexible two-parameter expo-power (EP) utility function which allows for varying degrees of relative risk aversion. The use of this function in experimental studies was popularized, in large part, by Holt and Laury (2002), who define the EP utility function as:

$$U(x) = \frac{(1 - \exp(-\alpha_{EP} x^{1-r_{EP}}))}{\alpha_{EP}} \tag{3}$$

The expected utility can then be evaluated in the manner described in Equation (2). Holt and Laury (2002) note that with this functional form, relative risk aversion is $r_{EP} + \alpha_{EP}(1 - r_{EP})x^{1-r_{EP}}$. Therefore, relative risk aversion varies with the size of the lottery prize x if $\alpha_{EP} \neq 0$. The EP function nests both constant absolute risk aversion (CARA) (as $r_{EP} \to 0$) and CRRA (as $\alpha_{EP} \to 0$), though it is not defined for $\alpha_{EP} = 0$.

Though EUT is an attractive and popular theory of decision-making under risk due to its relative simplicity, it does not seem to fit the empirical evidence well

[4] The Arrow-Pratt measure of relative risk aversion, or coefficient of relative risk aversion, is defined as $\frac{-xU''(x)}{U'(x)}$.

[5] In all the estimations in this chapter, the show-up fee (of 5 birr) in the experiment is excluded from the lottery prize x; this is common practice in empirical analyses of experimental data, and assumes that participants do not integrate their show-up fee with earnings from the experimental decision problems (Harrison and Rutström 2008). This also implies that in reference-dependent models (such as those involving loss aversion), the reference point does not include the show-up fee – for example, if the initial endowment or the payoff in the risk-free choice option in the decision problem is considered to be the reference point, these values do not include the show-up fee. However, the results remain substantively the same when the show-up fee is included in the measure of the lottery prize x.

– there is now a large body of literature, using both experimental and field data, which finds that decision-makers systematically violate its basic tenets and make decisions under risk which are not appropriately described by EUT (Tversky and Kahneman 1992, Humphrey and Verschoor 2004a, Tanaka et al. 2010). For example, numerous experimental studies have found that decision-makers evaluate gains and losses (relative to a reference point) in risky lotteries differently, and that the disutility of losses weigh more heavily than the utility of comparative gains – this phenomenon has been termed *loss aversion* (Kahneman and Tversky 1979, Tversky and Kahneman 1992). Further, Rabin (2000) notes that preferences incorporating loss aversion, unlike expected utility models, can explain both non-trivial risk aversion over small stakes and reasonable degrees of risk aversion over large stakes. One popular method to account for this particular violation of EUT is to utilize a two-part power utility (PU) function (see, for example, Tversky and Kahneman 1992, Harrison et al. 2007).

2.2 Loss aversion

The two-part power utility function, which is defined separately over gains and losses (relative to a reference point), has the following form:

$$U(x,\chi) = \begin{cases} (x-\chi)^{\alpha_P} & \text{if } (x-\chi) \geq 0 \\ -\lambda(-(x-\chi))^{\alpha_P} & \text{if } (x-\chi) < 0 \end{cases} \tag{4}$$

Thus, utility is no longer defined over the final prize amounts (x) in the lotteries, but is defined separately over gains and losses relative to a reference point χ.[6] Further, α_P is the parameter capturing the curvature of the utility function (referred to as the risk aversion parameter by Harrison and Rutström 2009) and λ is the loss aversion parameter. $\lambda > 1$ indicates loss aversion, and the magnitude of λ determines the extent to which losses loom larger than gains in decision-makers' valuation of risky lotteries (with higher values of λ indicating greater

[6] Koszegi and Rabin (2006, 2007) argue that the reference point is based on earnings expectations.

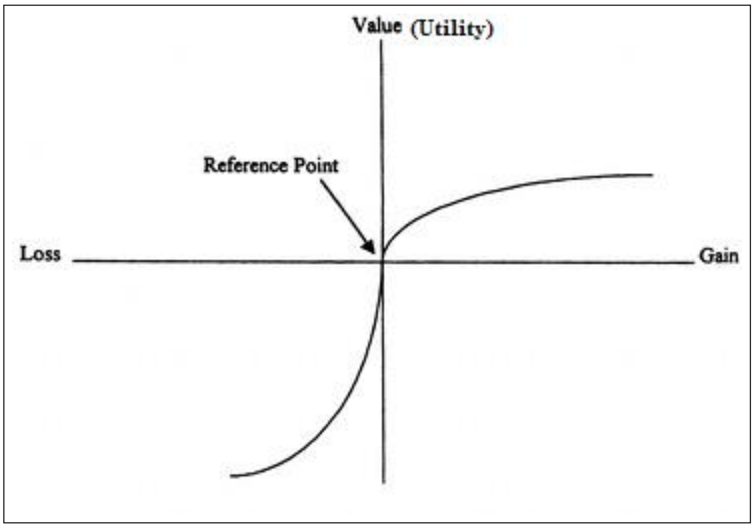
loss aversion). When $\lambda = 1$, there is no loss aversion, and gains and losses are evaluated equivalently. This form of utility was used by Tversky and Kahneman (1992) in their development of cumulative prospect theory, in order to shed light on loss aversion, stress the importance of reference points in evaluating lotteries, and to allow for the possibility that decision-making may be impacted by how lotteries (prospects) are framed (or presented to participants).

Tanaka et al. (2010) and Liu (2008), analyzing experimental data from rural Vietnam and China, respectively, find that most subjects in their experiments make choices inconsistent with EUT, and use a similar two-part power utility functional form with loss aversion to describe experimental choices. Additionally, numerous studies have found that individuals exhibit loss aversion in both experimental behaviour (for example, Humphrey and Verschoor 2004a, Barr and Packard 2005) as well as non-experimental behaviour (for example, Farber 2008, Crawford and Meng 2011). Indeed, Sprenger (2010) notes that such decision models of reference-dependent preferences have gained traction recently, and have been used to rationalize the endowment effect as well as important anomalies observed in labour market decisions, consumer behaviour and finance. Further, Fafchamps (2009) and Tanaka and Munro (2012) note that loss aversion may be a more important characterization of the behaviour of poor villagers in developing countries than risk aversion, as it plays a larger role in shaping individual preferences and driving decision-making.

Values of $\alpha_P < 1$ indicate a utility (or value) function which is concave above the reference point and convex below it; that is, $\alpha_P < 1$ reflects risk aversion for prize amounts greater than the reference point (gains), and risk seeking behaviour for prize amounts less than the reference point (losses) (Harrison and Rutström 2009).[7] This implies an “S-shaped” value function of the form illustrated in Figure 1. Further, loss aversion – indicated by values of $\lambda > 1$ – generates a kink at the reference point which makes the utility function steeper for losses than for gains around the reference point.

[7] $\alpha_P = 1$, on the other hand, reflects risk neutrality for gains and losses (Andersen et al. 2006a).

Figure 1. *S-shaped value (utility) function*



Source: Padula and Busacca (2005)

Using this utility function and weighting outcomes by objective probabilities, gain-loss utility from a lottery i (GLU_i) can be evaluated in much the same manner as described in Equation (2):

$$GLU_i = \sum_{k_i=1}^{K} p_{k_i} U_{k_i} \tag{5}$$

where p_{k_i} is the objective probability associated with outcome $k_i \in \{1, \ldots, K\}$ in lottery i.

Differential evaluation of losses and gains, however, is not the only deviation from EUT that has been observed in empirical studies. One of the most important deviations from EUT is considered to be the non-linear weighting of probabilities associated with final outcomes – Harless and Camerer (1994) observe that probabilities seem to enter decisions non-linearly, according to a *probability weighting function*, rather than linearly as assumed under EUT. Of the models which incorporate non-linear weighting of probabilities, rank-dependent utility theory has emerged as a leading theory of choice under risk (Harless and Camerer 1994, Humphrey and Verschoor 2004a).

2.3 Rank-dependent utility theory and cumulative prospect theory

Rank-dependent utility theory, developed by Quiggin (1982, 1993), does not assume that utilities from different outcomes in a lottery are weighted by objective probabilities, but instead allows decision-makers to use non-linear transformations of cumulative objective probabilities for weighting outcomes. RDU allows preferences to depend on the rank of the final outcome through probability weighting, and the probability weighting function can be interpreted as a decision weighting function which measures the subjective importance of different outcomes in the evaluation of a lottery (Humphrey and Verschoor 2004a, Harrison and Rutström 2008). Further, Humphrey and Verschoor (2004a) note that rank-dependent models represent a clear improvement on expected utility theory, and both rank dependence and loss aversion are important features of real-world behaviour under risk.[8]

There are two components in the RDU specification: the utility function and the probability weighting function. In the RDU framework, probability weighting can be combined with any of the three utility functions described earlier. However, instead of weighting outcomes by the objective probabilities p_{k_i}, RDU uses transformations of cumulative probabilities, w_{k_i}, to weight outcomes and evaluate lotteries, as follows:

$$RDU_i = \sum_{k_i=1}^{K} w_{k_i} U_{k_i} \tag{6}$$

where RDU_i is the rank-dependent utility from lottery i and the states (outcomes)

[8] Segal and Spivak (1990) define rank-dependent and loss aversion models as exhibiting first-order risk aversion, while expected utility models exhibit second-order risk aversion – first (second) order risk aversion implies that the risk premium for a small risk is proportional to the standard deviation (variance) of the risk. Further, Schmidt (1999) notes that the two types of risk aversion generate significantly different results, and models exhibiting first-order risk aversion can better explain real-world data and phenomena (for example, the equity premium puzzle).

in lottery i are ranked from worst (U_{1_i}) to best (U_{K_i}). w_{k_i} is defined as follows:

$$w_{k_i} = \begin{cases} \omega(p_{1_i}) & \text{for } k_i = 1 \\ \omega(p_{1_i} + \ldots + p_{k_i}) - \omega(p_{1_i} + \ldots + p_{(k-1)_i}) & \text{for } k_i > 1. \end{cases} \tag{7}$$

where $\omega(p)$ is some probability weighting function.[9]

As is clear from Equations (6) and (7), the choice of the probability weighting function (ω) is very important for RDU specifications. I consider two single-parameter functional forms for the probability weighting function ω that have been widely used in the experimental literature – the first proposed by Tversky and Kahneman (1992) and the second by Prelec (1998). Both have well-defined endpoints, with $\omega(p) = 0$ for $p = 0$ and $\omega(p) = 1$ for $p = 1$, and

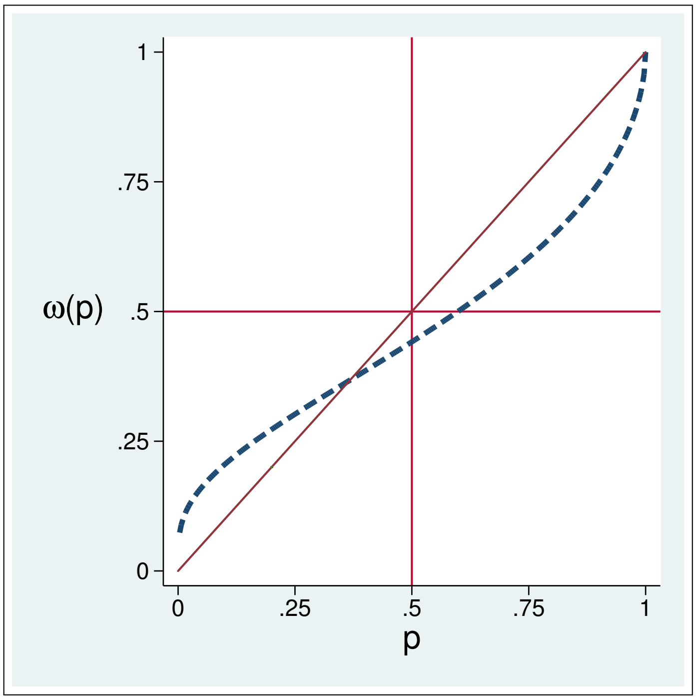
$$\text{Tversky-Kahneman function: } \omega^{TK}(p) = p^{\gamma}/[p^{\gamma} + (1-p)^{\gamma}]^{1/\gamma} \tag{8}$$

$$\text{Prelec function: } \omega^{P}(p) = \exp\{-(-\ln(p))^{\phi}\} \tag{9}$$

for $0 < p < 1$. Weighting of cumulative objective probabilities is implied by $\gamma \neq 1$ and $\phi \neq 1$. For standard experimental samples of university students in developed countries, it is commonly found that probability weighting follows an "inverse S-shape", concave for low probabilities and convex for high probabilities (Gonzalez and Wu 1999, Andreoni and Sprenger 2010). This corresponds to $0 < \gamma, \phi < 1$ for equations (8) and (9), and implies a systematic overweighting of small probabilities and an underweighting of large probabilities, as illustrated in Figure 2. Since probability weights can be interpreted as decision weights in the RDU setting, these parameter values indicate overweighting of extreme (unlikely)

[9] In the original version of prospect theory proposed by Kahneman and Tversky (1979), the probability weighting function was defined over individual rather than cumulative probabilities. However, Tversky and Kahneman (1992) note that this weighting scheme does not always satisfy stochastic dominance – Kahneman and Tversky's (1979) prospect theory does not ensure (or predict) that first-order stochastically dominated prospects are never chosen, and thus they are required to assume that transparently dominated prospects are eliminated in the editing phase. Further, their model is not readily extended to lotteries with a large number of outcomes. Since weighting over cumulative probabilities solves both these problems (by changing the preference for first-order stochastically dominant prospects), I only consider cumulative probability weighting, which is embodied in RDU (Tversky and Kahneman 1992).

Figure 2. *Inverse S-shaped probability weighting*



outcomes (Harrison and Rutström 2008). The determination of these probability weighting parameters and the identification of the general shape of the probability weighting function have received considerable attention in experiments as well as theory (Andreoni and Sprenger 2010).

When probability weighting is combined with the CRRA or EP utility function (providing the RDU-CRRA or RDU-EP specification, respectively), EUT is a special case of RDU (that is, when $\gamma = 1$ and $\phi = 1$) and the RDU model nests EUT. However, rank-dependent utility theory can also be extended to utility functions which are defined separately over gains and losses relative to a reference point (rather than final outcomes), such as the two-part power utility function described in Equation 4. This describes a class of decision models that comprise cumulative prospect theory (CPT); further, the two-part power utility and probability weighting specifications considered here – the combinations of which I refer to as CPT specifications throughout this thesis – are commonly used in studies involving CPT (for example, Harrison and Rutström 2009, Tanaka et al. 2010). CPT was developed by Tversky and Kahneman (1992) and has three critical features, as noted by Harrison and Rutström (2009): (i) the arguments of the utility function are gains and losses relative to some reference point, rather than the final outcome values in the decision problems; (ii) there is allowance for loss aversion, that is, for the possibility that losses loom larger than gains in the utility function; and (iii) there is allowance for the weighting of cumulative objective probabilities. Thus, CPT incorporates both sign- and rank-dependent preferences.[10] Starmer (2000), after reviewing the major experimental and non-experimental evidence, concludes that CPT is probably the strongest contender to EUT as a theory of decision-making under risk.

3 Outline of Empirical Investigation

In the previous section, I described some of the important models of decision-making under risk. In order to investigate which models best describe the behaviour under risk of participants in the experiment, I empirically analyze the choices made by participants in the experimental decision problems, within the framework of these models.

I begin by first estimating the CRRA utility function within the EUT framework, as described in Equations (1) and (2), using the choice data from the experimental decision problems. Assuming that the choices of participants are generated by this decision model, I aim to estimate the coefficient of relative risk aversion.[11] In order to account for, and determine, the extent to which participants make errors in the decision-making process, I extend the EUT-CRRA specification and aim to estimate it in conjunction with two different error models, the Luce and Fechner models. Staying within the framework of EUT, I also aim to estimate

[10] Since the RDU-CRRA and RDU-EP specifications defined here do not incorporate features (i) and (ii), which are critical aspects of cumulative prospect theory, I do not refer to them as CPT specifications.

[11] The importance of estimating this risk aversion parameter is highlighted by the work of Hill et al. (2011), who analyze data from the ERHS and find that the estimated coefficient of relative risk aversion in the EUT-CRRA specification plays a significant role in crucial insurance purchase decisions.

the more flexible two-parameter EP function, both with and without the error specifications. Using this function, which allows for varying degrees of relative risk aversion, I can analyze whether the assumption of CRRA is appropriate for explaining the decision-making behaviour under risk of the experimental sample.

The first departure from EUT is to allow the utility of participants to be defined over deviations from a reference point, and to allow for differing evaluations of gains and losses relative to that reference point. To do this, I estimate a two-part power utility function, which is defined separately over gains and losses (as detailed in Equation (4)); further, by estimating the loss aversion parameter, I can establish the extent to which participants in the experiment demonstrate aversion to losses. Additionally, unlike as done by most experimental studies utilizing similar reference-dependent decision models (for example, Liu 2008, Tanaka et al. 2010), I do not assume a particular exogenously-given reference point but instead estimate it, thus allowing the data to determine the reference point used by subjects in decision-making under risk (following the strategy of Andersen et al. 2006a, Harrison and Rutström 2008). While Tversky and Kahneman's (1992) original formulation of CPT includes the two-part power utility function in combination with probability weighting, I analyze the two-part power utility and CPT specifications separately, which facilitates the separate consideration of the significance of loss aversion and probability weighting.

In the analysis so far, I have assumed that participants' choices – the data in this case – are generated by a single model of behaviour under risk. However, Harrison et al. (2010) note that decision-making under risk is sufficiently heterogenous that it is unlikely to be described by a single theory or model; thus, in order to allow for the possibility that heterogenous theories of decision-making may co-exist in the same sample, I estimate a *mixture model* of the form specified by Harrison and Rutström (2009) and Harrison et al. (2010). Using the mixture model approach in combination with the two-part power utility and EUT specifications, I aim to evaluate the fraction of choices in the sample that are better characterized by each of these decision models.

Next, I utilize a RDU framework to explore whether there is significant probability weighting, that is, whether probabilities enter decisions linearly (objective probability weighting) or non-linearly (through a probability weighting function). I combine the two probability weighting functions (described in Equations (8) and (9)) and the cumulative weighting procedure (described in Equations (6) and (7)) with the CRRA, EP and two-part power utility functions; I aim to estimate the probability weighting parameter as well as the parameters of the utility functions. The estimate of the weighting parameter would indicate whether there is significant probability weighting, and also provide information on the shape of the probability weighting function; this would determine whether non-linear weighting of cumulative probabilities is important in explaining deviations from EUT (Humphrey and Verschoor 2004a). Thus, these estimations would help shed light on the suitability of rank-dependent utility theory and cumulative prospect theory as the latent decision-making process generating subjects' choices under risk. In order to estimate the parameters in the different models of decision-making under risk, I do not limit myself to a single estimation methodology, but use different estimators that impose varying degrees of structure (for example, the interval regression estimator and the structural maximum likelihood estimator, described by Harrison and Rutström 2008).

Finally, I investigate two interesting patterns observed in the data in the context of the empirical findings – the high level of insurance purchase in the index insurance decision problems and the systematic difference in choices in two numerically-identical (but differently-framed) decision problems of the experiment.

4 EXPERIMENT DESIGN

The experiment, conducted in November and December 2009, involved 378 subjects from seven sites of the Ethiopian Rural Household Survey (ERHS), spanning three regions of the country.[12,13] The experiment is described below, and this description is based on that provided by Clarke and Kalani (2011).

The experiment was designed to evaluate the demand for different types of formal insurance arrangements and to elicit risk preferences. It included a benchmark decision problem, framed in the abstract, and four framed insurance decision problems. In each session subjects played the benchmark decision problem and two of the four insurance decision problems. Therefore, subjects made three decisions and at the end of the session played, and were paid for, only one of the three decision problems; in addition, all participants were paid a show-up fee of 5 birr. Each subject randomly selected the problem (of the three he responded to) he would play for real money by choosing one out of three numbered tokens placed face down on a table. Harrison et al. (2010) and Charness et al. (2013) note that this procedure motivates participants to consider each choice carefully as if it were for real money, and provides sufficient incentive for participants to reveal preferences truthfully (without evoking wealth effects). The daily wage for casual farm labour in the areas where the experiment was conducted was between 15 and 20 birr (1.2 to 1.6 USD). Minimum and maximum earnings in the experiment were 5 and 80 birr and mean realized earnings, including the show-up fee, was 40 birr. Thus, the mean realized earnings represents a significant two to three days' income from casual farm labour in the experimental sites, and thus there is expected to be sufficient financial incentive for participants to answer truthfully in the decision problems.[14]

Further, drawing on best practice and based on extensive piloting, decision

[12] The seven sites chosen for the experiment were Sirbana Godeti, Korodegaga, Indibir, Milki, Komargefia, Karafino and Bokafia. Further, the subjects were chosen from households in the ERHS. These sites are close in distance to the capital of Addis Ababa (and span three regions of the country), and within these sites, all households who participated in the ERHS were invited to participate in the experiment. Approximately 96% of invited households attended the experiment.

[13] The experiment was designed and conducted by Daniel Clarke (University of Oxford), with financial support from the International Labour Organization (ILO) under the Microinsurance Innovation Facility and logistical support from the Center for the Study of African Economies (CSAE) and the Ethiopian Development Research Institute (EDRI).

[14] Falco (2012), analyzing experimental data from Ghana, notes that average experimental earnings close to the daily income of participants is sufficient to provide high enough incentive for subjects to truthfully reveal risk preferences.

problems were designed to be easily understood by subjects, choices and payoffs were described orally with the help of visual aids, randomization devices were physical and generated salient probabilities using familiar mechanisms, session money was physical, and understanding was confirmed and tested throughout the session (Barr and Genicot 2008).[15]

4.1 Benchmark decision problem

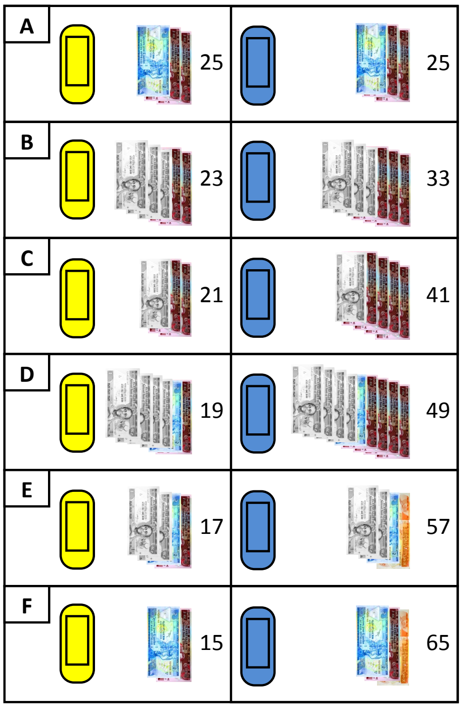
The benchmark decision problem (B) was as follows. Each subject was presented six lotteries, shown in the rows of Table 1, and was required to choose his most preferred option. Alternatives were ordered to be increasing in both the average payoff and the variance around that payoff. Alternative A was the "safe" option, offering a certain amount, and alternative F had the highest payoff mean and variance. Following Barr and Genicot (2008), the gamble was framed in the gain domain and, whichever gamble was chosen, the payoff was determined by playing a game that involved guessing which of the enumerator's hands contained a blue rather than a yellow counter.[16] The decision problem was explained privately to each subject, who made a private decision. After making decisions, subjects were seated separately and were not allowed to talk to each other.

In the terminology proposed by Harrison and List (2004), the benchmark decision problem represents an *artefactual field experiment*, that is, a conventional lab experiment framed in the abstract, but conducted with a non-standard subject pool of developing country inhabitants. All 378 participants in the experiment were presented with the benchmark decision problem, and the visual aid used for this decision problem is displayed in Figure 3.

[15] The use of physical randomization devices, rather than computers, helps mitigate the concern that subjects might not believe that the random process is actually fair (Harrison et al. 2010).

[16] In all the decision problems of this experiment, the blue counter (or token) represented the good individual outcome while the yellow counter represented the bad individual outcome.

Figure 3. *Presentation of the benchmark decision problem to experiment subjects*

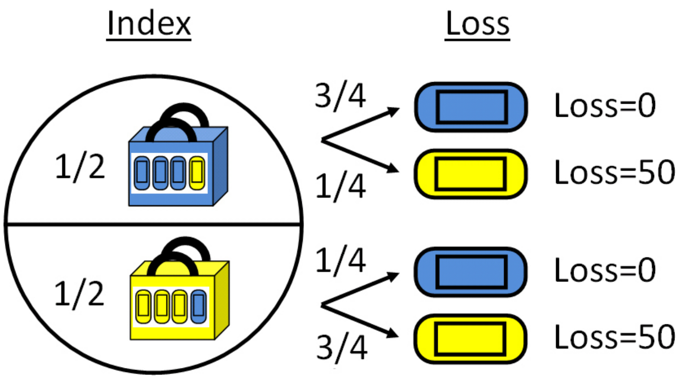


4.2 Insurance decision problems

Two of the four insurance problems involved indemnity insurance, while the other two involved index insurance. In an index insurance arrangement, the net transfer between insurer and policyholders depends only on a publicly verifiable index. For example, in the case of rainfall (or weather) indexed insurance, policyholders receive an insurance payout if the rainfall measured at a local weather station falls below a certain level. All four insurance purchase decisions were framed to be as similar as possible to real insurance purchase decisions, albeit in the controlled environment of the lab, with an objective probability structure, and with more time spent explaining and individually confirming understanding than would occur in the marketing process for a real product.

Additionally, all four insurance decision problems were framed in the loss domain. At the start of each insurance decision problem, each subject was

Figure 4. *Two-stage probabilistic structure for insurance decision problems*



physically given 65 birr of game money and told that he might lose 50 birr. Game money was smaller and more brightly coloured than Ethiopian currency but was otherwise recognizably similar. Additionally, enumerators spent 20 minutes explaining each insurance decision problem to the group of subjects, with an additional 10-20 minutes spent privately confirming understanding and recording decisions. Following common practice, enumerators referred to both index and indemnity insurance as insurance, rather than referring to the former as a derivative.

Subjects were randomly partitioned into pairs (whose role is explained below) and insurance purchase decisions had the following two-stage probabilistic structure (see Figure 4). First, a fair wheel was spun to determine whether a blue or yellow bag would be used for the pair of subjects. The blue bag contained three blue tokens and one yellow, and the yellow bag contained three yellow tokens and one blue. Second, each member of the pair chose one token from the selected bag, with replacement. An outcome for a pair therefore comprised a bag and two tokens.[17]

In addition to being given an explanation in terms of the wheel, bags and tokens, subjects were given an explanation for the probability structure in a real-world agricultural context. Subjects were told that the bags could be thought of as the weather, with the blue bag representing good weather and the yellow bag

[17] The experimental designs developed by Barr (2007) and Humphrey and Verschoor (2004b) also involve bags containing different coloured marbles or balls which reflect different outcomes.

representing poor weather. The tokens were likened to the actual yield on a plot, with a blue token representing a good year (in terms of yield) for the owner and a yellow token representing a bad year. Bad (good) weather was likely to lead to a bad (good) year for the owner, but this was not always the case: there was one yellow token in the blue bag and one blue token in the yellow bag.[18] In the terminology of Harrison and List (2004), the insurance decision problems represent *framed field experiments*, which are the same as *artefactual field experiments* but are framed in the context of actual decisions made in field, rather than in the abstract. Thus, the indemnity and index insurance problems are framed as real-world weather indemnity and weather indexed insurance purchase decisions.

Given this probability structure, the treatments may therefore be briefly summarized as follows. For individual treatments each subject was liable for his own loss in full: if a subject drew a yellow token he lost 50 birr. For group treatments the total loss for each pair was split evenly between the pair: each subject therefore lost 25 birr for each yellow token drawn by either member of the pair. For a particular subject, indemnity insurance then corresponded to purchasing insurance against himself (or his partner) drawing a yellow token and index insurance corresponded to purchasing insurance against his pair drawing a yellow bag. Indemnity insurance was priced with a loading of 60% and index insurance with a loading of 20%.[19]

The purpose of pairs was as follows. Each insurance decision problem was explained to all subjects in the session but subjects could only ask questions privately at the level of the pair. In individual treatments, the sole effect of

[18] Lybbert et al. (2010) conduct experiments in Peru and Kenya involving area yield indexed insurance that are framed in terms of agricultural yield and weather outcomes. They note that such experiments could help educate participants about relatively complex new products such as index insurance.

[19] An insurance loading is defined as (Premium Charged)/(Expected Claim Income) $-$ 1. Loadings of 20% and 60% are low compared to reported commercial loadings for crop insurance, which range from 70% to 430% for weather indexed insurance (Cole et al. 2009, Table 1) and 140% to 470% for indemnity insurance (Hazell 1992, Table 1). However, the 50% probability of claim payment in the experiment is much higher than that for commercial insurance products, and so these loadings cannot be directly compared.

Table 1. *Individual indemnity insurance purchase decision (T_{IM}) and benchmark decision (B)*

	Premium Choice (T_{IM})	Equivalent Choice in Benchmark (B)	Net Payoff (Ethiopian birr)		Expected Payoff	Risk Aversion Range (CRRA)*
Total Loss:			50	0		
Probability:			$1/2$	$1/2$		
	0	F	15	65	40	$(-\infty, 1.036)$
	8	E	17	57	37	$(1.036, 1.285)$
	16	D	19	49	34	$(1.285, 1.715)$
	24	C	21	41	31	$(1.715, 2.698)$
	32	B	23	33	28	$(2.698, 8.553)$
	40	A	25	25	25	$(8.553, +\infty)$

* Risk aversion range denotes range of coefficients for which choice would be optimal for a subject with CRRA preferences (Equation (1)) over earnings from the experiment, excluding show-up fee.

pairing was that each subject could hear any questions asked by his partner. In group treatments, a subject's earnings would also depend on his partner's random token draw. Pairs therefore did not have a strategic function; all insurance problems were individual decision problems, with own earnings depending only on own choices and chance. A more detailed description of the insurance decision problems is provided below.

Individual Indemnity (T_{IM}): In the individual indemnity decision problem T_{IM}, a subject incurred a 50 birr loss if a yellow token (bad individual outcome) was drawn, but could purchase between zero and five units of individual indemnity insurance against the loss occurring. One unit of indemnity insurance cost a premium of 8 birr and reduced the retained loss on drawing a yellow token by 10 birr. Each subject had to choose one out of six options, and could therefore pay 0, 8, 16, 24, 32 or 40 birr to reduce the maximum loss to 50, 40, 30, 20, 10 or 0 birr, respectively (see Table 1). The gamble choices available to individuals in T_{IM} were therefore numerically identical (with the same outcome values and associated probabilities) to those in B. However the framing of the choices was significantly different. 136 of the 378 participants in the experiment were involved in decision problem T_{IM}; the visual aids used for the four insurance problems are displayed in Figure 5.

Individual Index (T_{IX}): In the individual index insurance decision problem

Figure 5. *Presentation of the insurance decision problems to experiment subjects*

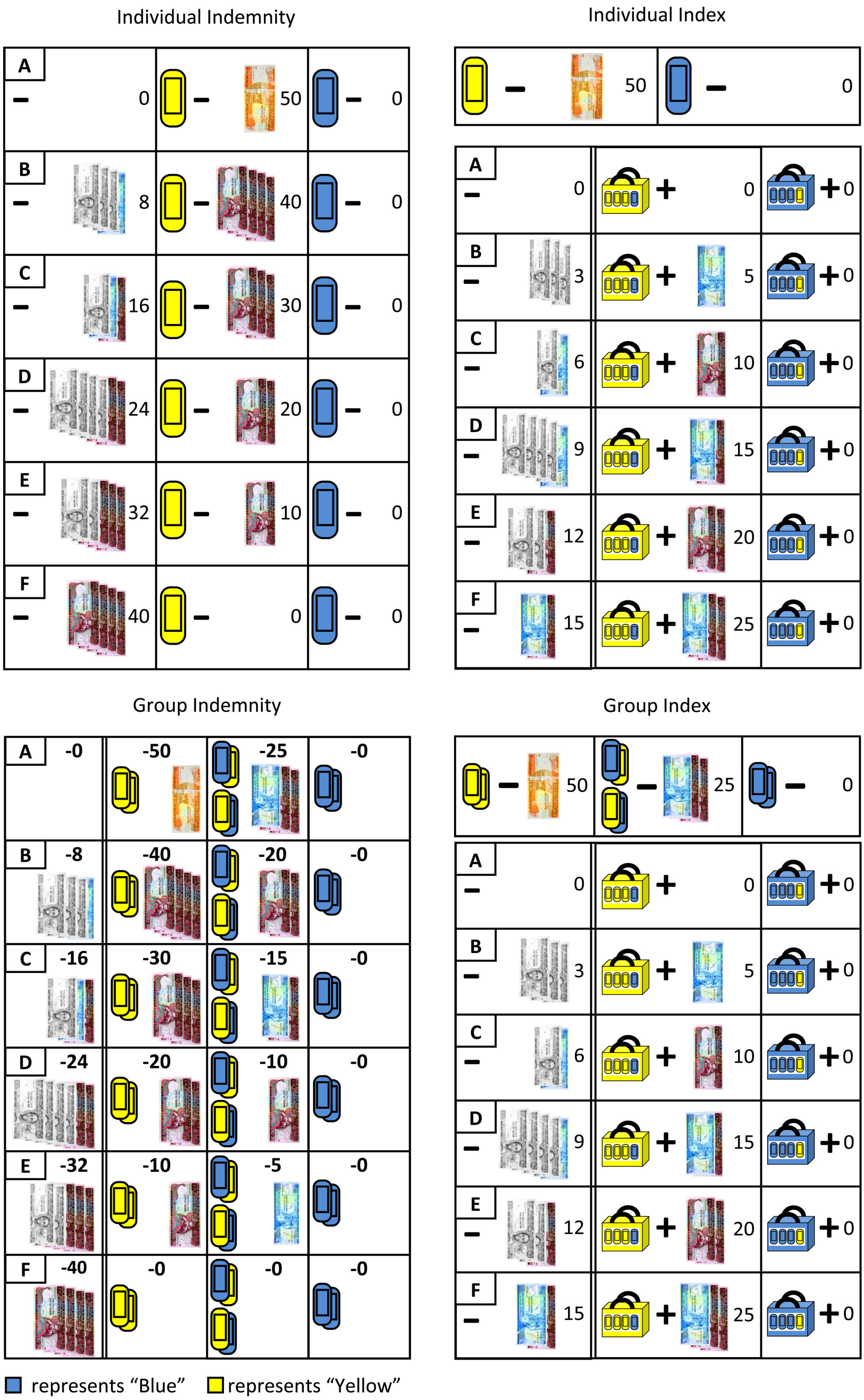


Table 2. *Individual index insurance purchase decision (T_{IX})*

	Premium Choice	Net Payoff (Ethiopian birr)				Expected Payoff	Risk Aversion Range (CRRA)*
Index:		Good	Bad	Good	Bad		
Loss:		50	50	0	0		
Probability:		$1/8$	$3/8$	$3/8$	$1/8$		
	0	15	15	65	65	40	$[0, 0.723)$ and $(3.888, +\infty)$
	3	12	17	62	67	39.5	$(0.723, 3.888)$
	6	9	19	59	69	39	N/A
	9	6	21	56	71	38.5	N/A
	12	3	23	53	73	38	N/A
	15	0	25	50	75	37.5	N/A

* Risk aversion range denotes range of coefficients for which choice would be optimal for a *risk averse* subject with CRRA preferences (Equation (1)) over earnings from the experiment, excluding show-up fee.

T_{IX}, instead of being able to insure against drawing a yellow token (crop loss), subjects could only purchase between zero and five units of index insurance against a yellow bag (bad index) being selected. One unit of index insurance cost a premium of 3 birr and led to a claim payment of 5 birr in the event of the yellow bag being selected, and zero otherwise (see Table 2). When describing T_{IX}, substantial emphasis was placed on the 1 in 8 chance of incurring a 50 birr crop loss (yellow token) despite the weather being good (blue bag) and therefore no claim payment being due (that is, the downside basis risk, or the risk that the income from index insurance will not accurately reflect the incurred loss). 258 of the 378 participants in the experiment were involved in decision problem T_{IX}.

Group Indemnity (T_{GM}): The group indemnity insurance decision problem T_{GM} was identical to T_{IM} except that, instead of losing 50 birr on drawing a yellow token, each subject lost 25 birr for each yellow token drawn by the pair, and one unit of indemnity insurance reduced the retained loss on drawing each yellow token by 5 birr. Each subject could therefore pay 0, 8, 16, 24, 32 or 40 birr to reduce the loss incurred from each yellow token draw to 25, 20, 15, 10, 5 or 0 birr, respectively (see Table 3). Subjects were told that group losses of 0 birr, 25 birr or 50 birr were approximately equally likely (the true probabilities were $10/32$, $12/32$ and $10/32$). Both members of a pair were subject to the

Table 3. *Group indemnity insurance purchase decision (T_{GM})*

	Premium Choice	Net Payoff (Ethiopian birr)			Expected Payoff	Risk Aversion Range (CRRA)*
Total Loss:		50	25	0		
Probability:		$5/16$	$6/16$	$5/16$		
	0	15	40	65	40	$(-\infty, 1.517)$
	8	17	37	57	37	$(1.517, 1.911)$
	16	19	34	49	34	$(1.911, 2.598)$
	24	21	31	41	31	$(2.598, 4.182)$
	32	23	28	33	28	$(4.182, 13.714)$
	40	25	25	25	25	$(13.714, +\infty)$

* Risk aversion range denotes range of coefficients for which choice would be optimal for a subject with CRRA preferences (Equation (1)) over earnings from the experiment, excluding show-up fee.

same uninsured crop loss for the pair but may have purchased different levels of insurance and therefore earned different amounts. 120 of the 378 participants in the experiment were involved in decision problem T_{GM}.

Group Index (T_{GX}): The group index insurance decision problem T_{GX} was identical to T_{IX} except that, instead of losing 50 birr on drawing a yellow token, each subject lost 25 birr for each yellow token drawn by the pair (see Table 4). Prices for, and payment from, index insurance were the same as in T_{IX}. Both members of a pair were subject to the same uninsured crop loss but may have purchased different levels of index insurance and therefore earned different amounts. 242 of the 378 participants in the experiment were involved in decision problem T_{GX}.

Note that in all decision problems, subjects were not given the option of indicating indifference between alternatives, and had to choose one of the six options. In the benchmark problem, option A had the least variance in outcomes while in the insurance problems option A corresponded to the least amount of insurance purchase. Further, there were three types of sessions, one with both individual insurance problems, one with both group insurance problems and one with both index insurance problems. Each set of three problems was presented in two different orders (see Table 5). Table 6 presents the frequency distributions of choices in the five decision problems.

Table 4. *Group index insurance purchase decision (T_{GX})*

	Premium Choice	Net Payoff (Ethiopian birr)						Expected Payoff	Risk Aversion Range (CRRA)*
Index:		Good	Bad	Good	Bad	Good	Bad		
Total Loss:		50	50	25	25	0	0		
Probability:		1/32	9/32	6/32	6/32	9/32	1/32		
	0	15	15	40	40	65	65	40	$[0, 0.566)$ and $(19.929, +\infty)$
	3	12	17	37	42	62	67	39.5	$(0.566, 0.688)$ and $(6.852, 19.929)$
	6	9	19	34	44	59	69	39	$(0.688, 0.771)$ and $(3.919, 6.852)$
	9	6	21	31	46	56	71	38.5	$(0.771, 1.000)$ and $(2.399, 3.919)$
	12	3	23	28	48	53	73	38	$(1.000, 2.399)$
	15	0	25	25	50	50	75	37.5	N/A

* Risk aversion range denotes range of coefficients for which choice would be optimal for a *risk averse* subject with CRRA preferences (Equation (1)) over earnings from the experiment, excluding show-up fee.

Table 5. *Decision problems*

Session type	Decision Problem			Number of sessions	Number of subjects
	First	Second	Third		
Individual	B	T_{IX}	T_{IM}	7	68
Individual (reversed)	T_{IM}	T_{IX}	B	7	68
Group	B	T_{GX}	T_{GM}	5	48
Group (reversed)	T_{GM}	T_{GX}	B	8	72
Index	T_{IX}	T_{GX}	B	6	58
Index (reversed)	T_{GX}	T_{IX}	B	6	64
				39	378

Table 6. *Frequency distributions of choices in the decision problems*

Choice	Decision Problem				
	B	T_{IM}	T_{IX}	T_{GM}	T_{GX}
A	7.14%	5.15%	3.10%	4.17%	2.89%
B	7.67%	18.38%	12.79%	15.83%	11.16%
C	11.90%	25.00%	17.83%	25.00%	14.88%
D	21.16%	14.71%	21.32%	17.50%	23.14%
E	20.37%	16.91%	19.77%	21.67%	19.42%
F	31.75%	19.85%	25.19%	15.83%	28.51%
Total Choices	378	136	258	120	242

The decision problems of the Ethiopian experiment use the Ordered Lottery Selection design developed by Binswanger (1980, 1981) – in which participants make a single choice out of multiple options (generally four to six) – to elicit risk preferences. Other popular implementations of this design are by Eckel and Grossman (2002) and Eckel and Grossman (2008). Whilst alternative methodologies (for example, the Multiple Price List procedure popularized by Holt and Laury 2002) have become popular in recent years for experiments with standard samples of university students in developed countries (Harrison and Rutström 2008), the simplicity of the Ordered Lottery Selection design makes it well suited for non-standard samples with low levels of formal education; thus, it has been used extensively in developing country settings with poorly-educated subjects (for example, Barr and Genicot 2008, Van Campenhout et al. 2008, Attanasio et al. 2012).[20,21] In addition, Harrison and Rutström (2008) observe that this experimental design has transparent incentives for truthful responses and is very portable in the field.

It should be pointed out that the benchmark decision problem B differed from the insurance decision problems in three important ways. First, B was framed in the gain domain while the insurance problems were framed in the loss domain with the initial endowment of 65 birr provided up-front at the start of the decision problem.[22] Harbaugh et al. (2002) note that there could be significant framing effects that cause subjects to respond to decision problems framed in the gain and loss domains differently. Second, the insurance problems involved decisions about insurance purchase and premium payments, while B was a simple Binswanger (1980) lottery framed in the abstract without any insurance aspect. This, too, may be the cause of framing effects which could lead to systematically different choices in the insurance decision problems, as compared to the choices in B. Third, the compound probability structure for insurance treatments was much more complex[23] than the simple fair draw for the benchmark. Spears

[20] Strictly speaking, in the experimental decision problem of Holt and Laury (2002), the payoffs remain the same in each of the ten binary questions while the probabilities differ. However, throughout this thesis, I refer to any decision problem that involves multiple binary choices (even those in which the payoffs differ in the binary questions) as utilizing the Holt and Laury (2002) experimental design, and use this term interchangeably with the Multiple Price List elicitation procedure. Similarly, even though the original experiments of Binswanger (1980) involved only probabilities of 0.5, I refer to any decision problem involving a single choice between more than two options as utilizing the Binswanger (1980) experimental design, and use this term interchangeably with the Ordered Lottery Selection elicitation procedure and Binswanger-style lottery. This is because these two elicitation procedures were popularized by, and are synonymous with, the above-mentioned experimental studies. Charness et al. (2013) note that this is common practice in the experimental literature (for example, Dave et al. 2010), and this classification aids the comparison of the two procedures by highlighting the main differences between them.

[21] Additionally, the number of choices elicited per subject in the Ethiopian experiment (three) is in line with the number elicited in other experiments using Ordered Lottery Selection problems (for example, Barr 2003a, Eckel and Grossman 2008, Dave et al. 2010).

[22] In the experiment conducted by Barr and Packard (2005) in Peru, the insurance decision problem was also framed in the loss domain.

[23] Whilst experimental economists might argue that this probability structure could be too complex for subjects to understand, it is much less difficult to understand than the joint probability structure for a real weather indexed insurance policy; although farmers might have a good understanding of the marginal loss distribution for their farms, they are unlikely to have a good understanding of the conditional distribution of weather indexed claim payments (Giné et al. 2005, Giné et al. 2007, Hill and Nobles 2011). Moreover, each stage of the randomization device was chosen to have salient probabilities of $1/4$, $1/2$ and $3/4$, and understanding was confirmed and tested throughout the session.

(2012), analyzing experimental and survey data from El Salvador and India, finds evidence of aversion to compound risk – participants prefer simple lotteries (which depend on only a single random event) to compound lotteries (which depend on multiple, sequential random events), even when the simple lotteries provide lower expected value. Thus, we would *ceteris paribus* expect participants in the Ethiopian experiment to exhibit greater risk aversion in the insurance decision problems – which involved compound lotteries – than in the benchmark decision problem (which involved a simple lottery). In addition, Crosetto and Filippin (2013) find that elicited risk preferences vary significantly between decision problems of differing complexity, which may be due to the invoking of different processes of decision-making under risk.

These important differences in the framing and nature of the benchmark and insurance problems are likely to lead to *preference reversals* on the part of subjects, which cause a systematic change in decisions (and the decision-making process) in otherwise identical tasks.[24] Such preference reversals are commonly found in lab experiments conducted in both developed and developing countries, and much of the evidence on the topic indicates that elicited risk preferences are not stable across elicitation procedures, as well as the degree of complexity, context, domain and framing of decision problems (Humphrey and Verschoor 2004a, Crosetto and Filippin 2013, Charness et al. 2013). Given that all these factors could significantly impact, and cause to differ systematically, the decisions subjects make in the experimental problems, I separately analyze the decisions in B and those in the four insurance problems, thus allowing for the possibility that different decision-making processes generate the choices in the benchmark and insurance problems. This provides a dataset of 378 choices in decision problem B and 756 choices in the insurance problems.[25]

[24] This phenomenon of preference reversals is explored further in Section 7.
[25] The choices in the four insurance problems are considered jointly.

5 EMPIRICAL ANALYSIS

5.1 Expected utility theory

I begin the empirical analysis of the experimental data by estimating the CRRA utility function within the EUT framework, as described in Section 2.1. In order to do this, I use a structural model combined with a maximum likelihood estimation technique, following the estimation strategy pioneered by Camerer and Ho (1994), and detailed in the work of Harrison and Rutström (2008). I also use an interval regression model to estimate the coefficient of relative risk aversion r in the CRRA utility function. Both these methodologies have been widely used in analyses of experimental data.

The structural estimation methodology is as follows (a detailed derivation of the structural model, including the specific assumptions underlying it, is provided in Appendix A). First, the expected utility EU_i^d from each potential lottery choice $i \in \{1, \ldots, 6\}$[26] in each decision problem $d \in \{B, T_{IM}, T_{IX}, T_{GM}, T_{GX}\}$ is calculated according to Equation (2), assuming that the utility of income is defined by the CRRA function in Equation (1).[27] The latent index ∇EU_i^d is then calculated as follows:

$$eu_i^d = \exp(EU_i^d) \tag{10}$$

$$\nabla EU_i^d = \frac{eu_i^d}{\sum\limits_{j=1}^{6} eu_j^d} \tag{11}$$

The latent index ∇EU_i^d, based on latent preferences, is in the form of a probability between 0 and 1, and thus can be directly linked to the observed choices. Following Harrison and Rutström (2008), ∇EU_i^d is interpreted as the probability of a subject choosing lottery choice i in decision problem d, and is used to construct the log-likelihood function. Thus, given the observed choice

[26] The choices A through F in the decision problems are referred to as choices 1 through 6 in this model description.

[27] The lottery prizes – and corresponding probabilities – in the decision problems are given in Tables 1 to 4.

$y_a^d \in \{1, \ldots, 6\}$ of individual a in decision problem d, the log-likelihood of the observed responses, conditional on the EUT-CRRA specification representing the true decision-making process of subjects[28], is

$$\ln L^{EUT}(r; \boldsymbol{y}) = \sum_{a=1}^{378} \ln(\nabla EU^{B}_{y_a^B}) \tag{12}$$

for decision problem B, and

$$\begin{aligned} \ln L^{EUT}(r; \boldsymbol{y}) \quad = \quad & \sum_{a=1}^{136} \ln(\nabla EU^{T_{IM}}_{y_a^{T_{IM}}}) + \sum_{a=1}^{258} \ln(\nabla EU^{T_{IX}}_{y_a^{T_{IX}}}) \\ & + \sum_{a=1}^{120} \ln(\nabla EU^{T_{GM}}_{y_a^{T_{GM}}}) + \sum_{a=1}^{242} \ln(\nabla EU^{T_{GX}}_{y_a^{T_{GX}}}) \end{aligned} \tag{13}$$

for the four insurance decision problems – T_{IM},T_{IX},T_{GM}, and T_{GX} – whose choices are considered jointly. Thus, there are 378 choices incorporated in the estimation for decision problem B and 756 choices incorporated in the estimation for the insurance problems ($\boldsymbol{y}$ denotes the vector of observed choices incorporated in the estimation). The log-likelihood function $\ln L^{EUT}$ can then be maximized with respect to the core parameter r to yield the maximum likelihood estimate of the coefficient of relative risk aversion r. The results of these maximum likelihood estimations are presented in Table 7.[29,30] This structural maximum likelihood estimation procedure, including the latent index described in Equation

[28] Note that here I assume that all the choices in the dataset are independent, and generated by a single model of decision-making under risk (in this case, EUT with CRRA utility), following the strategy of Harrison et al. (2007) and Harrison and Rutström (2008).

[29] The maximum likelihood estimations reported in this chapter were carried out in Stata – using the Newton-Raphson and Berndt-Hall-Hall-Hausman optimization algorithms – and MATLAB (using the sequential quadratic programming optimization algorithm).

[30] It is important to note that these estimations, along with all other estimations in Section 5, involve a between-subject analysis (as the choices of different subjects are combined in the estimations) based on functional form assumptions regarding the decision-making process of participants, following the strategy of Harrison et al. (2007) and Harrison and Rutström (2008). While this strategy is somewhat restrictive and relies on functional form assumptions, it is a constraint imposed by the data – there are too few answers elicited per subject (only three) in the Ethiopian experiment to conduct more flexible within-subject analyses (as done by Liu 2008, Tanaka et al. 2010) or non-parametric analyses (as done by Hey and Orme 1994, Holt and Laury 2002). Indeed, Harrison and Rutström (2008) note that within-subject analyses require a very large sample of choices for each participant (for example, the choices of each subject in a number of Multiple Price List decision problems). However, in order to mitigate the possibility of obtaining misleading results due to the selection of an incorrect functional form, I consider a wide range of decision models which incorporate different preference parameters.

Table 7. *Structural maximum likelihood estimates of risk attitudes for EUT with CRRA utility*

Coefficient	Estimate	Standard Error	p-value	Lower 95% Confidence Interval	Upper 95% Confidence Interval
A. Benchmark Decision Problem					
r	0.537	0.025	0.000	0.488	0.586
B. Insurance Decision Problems					
r	0.774	0.023	0.000	0.729	0.818

Wald test p-values and 95% confidence intervals reported.
The log-likelihoods corresponding to the estimations in Panels A and B are -629.4 and -1408.9.

(11), is also used by Harrison and Rutström (2008) to analyze the experimental choices of university students in Binswanger-style lotteries.

The coefficient of relative risk aversion is estimated to be 0.537 for the benchmark decision problem and 0.774 for the insurance decision problems; using Wald tests, I find that both estimates are statistically significantly different from 0 at the 1% level.[31,32] The structural estimation methodology used to estimate model parameters makes the traditional representative agent assumption that each subject in the experiment has the same true preferences; that is, each participant's choices under risk are assumed to be generated (with error) by the

[31] I also conducted likelihood ratio tests in addition to Wald tests to test the various hypotheses in this chapter – in all cases the conclusions from both tests are identical. Additionally, there are no notable changes in the results when standard errors clustered at the session level and robust to heteroskedasticity are used, indicating that heteroskedasticity and correlation of the error term between individuals in a particular session may not be major concerns.

[32] It is important to note this structural model incorporates the possibility of stochastic errors in decision-making, and thus represents a model of "stochastic choice with deterministic preferences" – the errors are modeled and incorporated using the logit discrete choice model developed by McFadden (1974). However, in this model, the extent (or magnitude) of these errors is fixed (as done by Harrison and Rutström 2009, Harrison et al. 2010). When preference parameters are assumed to be constant across individuals, these logit errors account for heterogeneity in the observed choices of participants. Appendix A provides more details on this structural model – in the context of this model, the assumption of a fixed extent of errors implies assuming the logit error term to have a fixed variance (and thus the scale of utility is normalized). The Luce and Fechner error models analyzed in Section 5.2, on the other hand, allow this variance (and thus the magnitude of noise in decision-making) to vary. There is substantial evidence of errors in decision-making in the experimental literature, and thus it crucial to allow for errors in the decision-making process of subjects; thus, any model used to describe experimental data (that is, the data-generating process) should be specified stochastically (Starmer 1999, Humphrey and Verschoor 2004a).

same decision model and the same preference parameter values (see Appendix A). This assumption has been incorporated in similar estimations by Andersen et al. (2006a), Harrison and Rutström (2009) and Harrison et al. (2011). Thus, the parameter estimates of the decision models can be interpreted as the parameter values for the average subject. The estimates in Table 7 indicate that if EUT with CRRA utility is assumed to represent the true decision-making process, the participants in the experiment exhibit moderate risk aversion, on average. These estimates of the average r are comparable to the coefficient of relative risk aversion of around 0.5 recently estimated for Ethiopian, Ugandan and Indian subjects in a traditional laboratory experiment (framed in the abstract) conducted by Harrison et al. (2010); additionally, Botelho et al. (2005) estimate r to be approximately 0.6 when analyzing experimental data from Timor-Leste. The estimates are also similar to those obtained using laboratory experiments in developed countries and comparable statistical methods (Harrison et al. 2010).

Additionally, the estimated coefficient of relative risk aversion is higher in the insurance problems than in the benchmark problem – using Wald and likelihood ratio tests, I can reject, at the 1% significance level, that the two coefficients are equal. This implies that if EUT-CRRA represents the true decision-making process generating choices in the problems, the risk aversion implied by the choices in the insurance problems is significantly higher than that implied by the choices in the benchmark problem. In other words, the statistically significant difference between the two coefficients provides evidence that participants evaluate the benchmark and insurance decision problems differently and exhibit significantly greater risk aversion in the insurance problems (possibly due to aversion to compound risk, as indicated by Spears 2012); it also highlights the importance of separately considering the benchmark and insurance choices in the estimations. Alternatively, it could indicate that different decision-making processes generate the choices of participants in the benchmark and insurance problems, possibly due to differences in the framing and complexity of these problems – this, including the phenomenon of preference reversals, is explored in later sections. Further, Barr (2007) find that in an experiment conducted in Peru, the coefficient of relative risk aversion elicited using a decision problem in the gain frame is more significantly correlated with real investment decisions, while that elicited using a decision problem framed in the loss domain is more significantly correlated with real insurance purchase decisions.

I now turn my attention to the interval regression model. If subjects are assumed to make decisions in accordance with the EUT-CRRA decision model (given by Equations (1) and (2)), then each choice in the problems B, T_{IM} and T_{GM} corresponds to a particular range of the coefficient of relative risk aversion r (see Tables 1 and 3). This is the range of coefficients for which that particular choice would be optimal for an individual with EUT-CRRA preferences. The interval regression model uses information on these bounds of risk aversion implied by the observed choices in the decision problems (Harrison and Rutström 2008), and has been utilized by Wik et al. (2004), Lusk and Coble (2008) and Tanaka and Munro (2012) in the context of risk preference measurement.[33]

There are certain choices (reflecting high amounts of insurance purchase) in the index insurance decision problems, T_{IX} and T_{GX}, which are not optimal for any expected utility maximizer with preferences satisfying risk aversion and constant relative risk aversion (that is, for any value of $r > 0$) (see Tables 2 and 4).[34] This is because if the individual is averse enough to risk to want to purchase the cover, he would also be averse enough to the downside basis risk associated with

[33] More details on the interval regression model are provided in Chapter 2, as well as in Harrison and Rutström (2008).

[34] Note that this upper bound on optimal insurance purchase is a consequence of the theoretical model of Clarke (2011). Though this theory applies only to risk averse individuals, it is highly likely that the entire Ethiopian sample – comprising of experiment participants who are relatively poor – falls within this category. The structural maximum likelihood estimations of the EUT-CRRA model also provide some evidence of this. See Section 6 for more details.

Table 8. *Interval regression estimates of risk attitudes for EUT with CRRA utility*

Coefficient	Estimate	Standard Error	p-value	Lower 95% Confidence Interval	Upper 95% Confidence Interval
A. Benchmark Decision Problem					
r	1.469	0.167	0.000	1.141	1.797
B. Indemnity Insurance Decision Problems					
r	4.547	0.289	0.000	3.980	5.114

Wald test p-values and 95% confidence intervals reported.
The log-likelihoods corresponding to the estimations in Panels A and B are -841.6 and -630.3.

index insurance to limit the size of the hedge.[35] Therefore, this model cannot be used to estimate r from choices in the two index insurance decision problems, since it does not allow for errors that cause observed choices to deviate from true preferences (as incorporated in the structural model described earlier); as a result, I only use the interval regression model to estimate r from the choices in the benchmark decision problem and (separately) from the choices in the two indemnity insurance problems, T_{IM} and T_{GM} (which do not involve basis risk).[36] The results of the interval regression analysis of choices in decision problem B (378 choices) and of choices in decision problems T_{IM} and T_{GM} (256 choices) are presented in Table 8.

[35] However, I still find that a considerable number of subjects chose options in the index insurance problems which are inconsistent with expected utility theory and (risk averse) constant relative risk aversion, indicating a direct violation of this decision model for these problems (if stochastic errors in decision-making are not considered). In line with the work of other experimental studies (for example, Harrison and Rutström 2008, Harrison et al. 2010), I begin the empirical analysis in this section assuming EUT with CRRA utility as the true decision-making process, and then go on to explore the validity of EUT as compared to alternative models, such as RDU and CPT, in later sections. Additionally, this issue related to high levels of index insurance purchase is further analyzed in Section 6.

[36] The interval regression model incorporates a maximum likelihood estimator and uses the implied range of r that each subject implicitly selected when making a particular choice as the latent variable or the implicit dependent variable. However, while Harrison and Rutström (2008) allow r to be a linear function of observable individual characteristics, I restrict r to be a constant, since I only aim to estimate the average r of the sample, in line with the representative agent assumption in the structural model; this enables me to compare the estimates from the interval regression and structural estimations. In Chapter 2, where I analyze the determinants of risk preferences exhibited in the Ethiopian experiment, I conduct estimations that assume r to be a linear function of observable characteristics.

The interval regression estimates indicate that the average coefficient of relative risk aversion r is 1.469 for the benchmark decision problem and 4.547 for the two indemnity insurance decision problems. Once again, the estimates of r are positive and statistically significant, indicating that the choices in these problems reflect risk aversion on the part of subjects. The estimates are also considerably greater than those obtained using the structural estimation method, reflecting a relatively higher level of risk aversion – the estimate of r from the indemnity insurance problems is particularly large.

The substantial difference in the estimates from the two models indicates that the distinct parametric assumptions (for example, the differing assumptions regarding errors in decision-making) made by the models matter significantly for the estimation of risk attitudes; this has been noted by Harrison and Rutström (2008) as well. While the structural model described earlier incorporates stochastic errors in decision-making, the interval regression model does not allow for errors that cause observed choices to deviate from true preferences. However, as noted by Humphrey and Verschoor (2004a), models of decision-making under risk should be specified stochastically, given that discrete choice is highly stochastic and numerous studies have found evidence of significant errors in experimental decision-making. As a result, it is likely that the structural maximum likelihood estimates provide a better indication of the true risk preferences of participants than the interval regression estimates.

To put the interval regression estimates into perspective, it is important to bear in mind that Harrison and Rutström (2008) survey a number of important experimental studies and find that most obtain an average r estimate between 0 and 1. However, once again, the coefficient estimates in Table 8 indicate that participants exhibit significantly more risk aversion in the indemnity insurance problems than in the benchmark problem, assuming that EUT-CRRA represents the true decision-making process for choices in these decision problems.

It is also important to note that the interval regression model is only suitable for analyzing choices using single-parameter decision models, such as EUT-CRRA, and therefore cannot be used for the estimation of non-EUT models considered in later sections, such as RDU and CPT, which incorporate multiple parameters.

5.2 Stochastic errors in experimental decision-making

Harrison and Rutström (2008) note that an important extension of the core deterministic choice model is to allow for subjects to make errors in decision-making, as discrete choice is highly stochastic; if purely deterministic choice is assumed without any allowance for errors in decision-making, a single observation in violation of a particular decision theory would be sufficient to reject it (Starmer 1999). While the structural model used in the earlier section does incorporate errors which allow observed choices to differ from true preferences (see Appendix A), it assumes that the extent of errors in decision-making is fixed. However, in this section I allow the magnitude of noise in decision-making to vary and attempt to estimate the parameter that determines this magnitude, thus shedding light on the extent of errors in decision-making. It is important to explicitly measure and account for the magnitude of noise in decision-making, since numerous studies have found evidence of substantial errors in the decision-making process of experiment participants; in addition, the one-shot nature of the experiment and the unfamiliarity of the Ethiopian participants with such decision problems are likely to induce substantial noise in revealed preferences (Humphrey and Verschoor 2004a). Indeed, Hey (1995) and Loomes et al. (2002) find that "expected utility plus white noise" represents an important model, and a good approximation, of decision-making under risk.

Harrison and Rutström (2008) and Harrison et al. (2011) suggest two important "white noise" stochastic error models that have been extensively used in experimental studies – the Luce error model and the Fechner error model. These error specifications – combined with decision models (such as EUT, RDU and CPT) – represent models of "stochastic choice with deterministic preferences", since the true preferences are deterministic (and given by the corresponding decision model) but the observed choices are stochastic due to the presence of "white noise" error – this error allows observed choices to deviate from true preferences (Hey and Carbone 1995).[37] These stochastic error models place restrictions on the violations or errors made by participants, and provide useful statistical information and instructions (Wilcox 2008). Indeed, Wilcox (2008) stresses the importance of further study on such stochastic error models, as these error models can greatly affect inferences regarding patterns of risk preferences. I aim to estimate both error specifications in conjunction with the EUT-CRRA and EUT-EP decision models.

Luce error specification: In this structural error specification, originally due to Luce (1959) and popularized by Holt and Laury (2002), the expected utility EU_i^d for each potential choice i in decision problem d is calculated (as a function of r) using the CRRA utility function described in Equation (1) and the expected utility formulation in Equation (2) (see Appendix B for a detailed derivation of this error specification). Then, the latent index ∇EU_{iLuce}^d is calculated as follows:

$$eu_{iLuce}^d = \exp(EU_i^d/\mu) \tag{14}$$

$$\nabla EU_{iLuce}^d = \frac{eu_{iLuce}^d}{\sum\limits_{j=1}^{6} eu_{jLuce}^d} \tag{15}$$

where μ is the Luce error parameter.[38] As $\mu \to 0$, the specification collapses to the deterministic EUT model, where the choice is strictly determined by the expected utility of the different lotteries, and the option with the highest expected

[37] Throughout this thesis, I refer to a deterministic preference-maximizing model with an added error structure as a model of "stochastic choice with deterministic preferences".

[38] Following Harrison and Rutström (2008) and Harrison et al. (2011), I refer to this parameter as the "error" or "noise" parameter, since it determines the extent to which choice probabilities depart from true preferences (given by expected utility theory) and thus the magnitude of stochastic errors in experimental decision-making. Additionally, following the classification of Harrison and Rutström (2008), I distinguish the "error" models described in this section from the models described in the previous section, even though both incorporate stochastic choice. See Appendix B for more details on these error models.

utility is always chosen (Lee 2008, Harrison and Rutström 2008). Larger values of μ imply greater errors in the decision-making process and a greater deviation of observed choices from true preferences (given by the expected utilities of the lottery choices), assuming that the true preferences of participants are given by EUT-CRRA (in this case). As $\mu \rightarrow \infty$, participants' choices among the six options in each decision problem become random, and completely independent of true preferences. When $\mu = 1$, this specification collapses to the structural specification described in Section 5.1, and the latent index $\nabla EU^d_{iLuce} = \nabla EU^d_i$. Therefore, the structural model used in the previous section (which does not involve the free noise parameter μ) incorporates errors in decision-making (that is, μ is not assumed to be 0) and thus allows for stochastic choice, but fixes the extent of errors in decision-making by assuming a fixed noise parameter (that is, through the $\mu = 1$ normalization).[39] That is, both specifications assume that EUT-CRRA plus noise represents the true data-generating process, but the Luce error specification allows the magnitude of the error (noise) parameter to vary and be estimated.

∇EU^d_{iLuce} is once again interpreted as the probability of a subject choosing lottery choice i in decision problem d, and is used to construct the log-likelihood function for the Luce error model for both the benchmark and insurance decision problems, in the manner described in Equations (12) and (13). The log-likelihood can then, in theory, be maximized with respect to the two parameters, r and μ, to obtain maximum likelihood estimates of these parameters.

However, when I attempted to maximize this log-likelihood function, the optimization routine did not converge, both for the data from the benchmark problem and the insurance problems. To explore the cause of this non-

[39] As noted earlier, in reference to the models described in Appendix A and B, this implies that the variance of the logit error term (in the logit discrete choice model) is fixed and the scale of utility is normalized. Numerous studies in the experimental literature use structural specifications that assume $\mu = 1$, as they permit errors but do not increase the number of parameters to be estimated (for example, Harrison and Rutström 2009, Harrison et al. 2010). On the other hand, the Luce error model described in this section does not assume $\mu = 1$ – it follows the logit discrete choice framework but allows the variance of the logit error term (and thus the magnitude of noise in decision-making) to vary; the value of μ determines the size of this variance.

convergence, I constructed simulated datasets of choices in the decision problems (for a range of feasible values of r and μ), of the same size and form as the actual experimental dataset at hand, assuming that EUT-CRRA plus Luce "white noise" error represents the true data-generating process; I constructed separate simulated datasets for the benchmark and insurance problems (with the same lotteries and payoffs as in these problems). I then used the structural model and maximum likelihood technique, as described above, to estimate the two parameters (r and μ).

I find that both parameters are estimated using this simulated data. However, I find that while the maximum likelihood estimate of r is very close to the value used to create the simulated data – with a relatively narrow confidence interval – the estimate of μ is very different from the value used to construct the simulated data, and the standard error is quite large. Even when the size of the simulated dataset was increased to 10,000 choices, the estimated value of μ deviated significantly from the true value, with a very large standard error – this was the case for a wide range of different values of r and μ used to simulate the data. Only when the size of the simulated dataset was increased to the order of 100,000 choices was μ estimated with any notable accuracy. Thus, the model itself is identified, but I am unable to estimate both parameters using the actual experimental data at hand.[40]

The cause of this inability to jointly estimate r and μ is most likely the nature of the Ethiopian experimental data, and in particular, the elicitation procedure used in the experiment. In this experiment, participants chose one out of six lotteries (following the Ordered Lottery Selection design developed by Binswanger 1980) for each of the three decision problems they were assigned. Most studies which estimate the Luce error specification (as well as non-EUT decision models) use Multiple Price List (MPL) decision problems, in which participants choose their preferred option from a pair of prospects, for a series of such pairs. In

[40] Lee (2008), analyzing experimental data from the United Kingdom, also encounters convergence problems when estimating the Luce error model using relatively few observations.

this elicitation method, participants make a relatively large number of binary choices (generally 10-30) in each decision problem; for example in the decision problem of the popular Holt and Laury (2002) experiment, each of the 212 subjects made decisions over 10 lottery pairs. Thus, the low power of the Ordered Lottery Selection decision problems of the Ethiopian experiment for jointly estimating both parameters is likely due to the low “precision” and coarser characterization of risk preferences associated with this elicitation procedure (Crosetto and Filippin 2013) – while the Ordered Lottery Selection decision problem involves only a single choice and thus enables the classification of subjects into relatively few preference categories (in the case of the Ethiopian experiment, six), the Multiple Price List procedure involves a relatively large number of choices and thus enables the classification of subjects into a greater number of preference categories. Further, this also leads to a considerably smaller number of total choices in an Ordered Lottery Selection experiment dataset, for a similar number of decision problems and a similar number of participants. Indeed, Charness et al. (2013) note that the greater precision of the preference parameter intervals provided by complex procedures like the MPL increases the likelihood of obtaining accurate point estimates of preference parameters using econometric methods.[41]

Thus, the simulations indicate that the Luce error model itself is identified – given an extremely large sample of choices in decision problems (of the form in the Ethiopian experiment) with the Ordered Lottery Selection experimental design, both r and μ can be estimated accurately (Cameron and Trivedi 2005). However, eliciting the number of responses required for the accurate estimation of these parameters using Ordered Lottery Selection problems is not feasible; with a dataset size of the order of magnitude of the Ethiopian dataset, only r –

[41] However, the MPL procedure is cognitively more challenging (that is, it places a greater “cognitive load” on subjects), and existing evidence indicates that MPL decision problems are not well understood by subjects with low education and literacy levels, making this procedure less suitable for accurately eliciting risk preferences in developing country settings (as compared to simpler procedures, like the Ordered Lottery selection method) (Dave et al. 2010, Charness and Viceisza 2012). This issue is explored in Chapter 3 using MPL data.

and not μ – can be estimated accurately. Thus, the parameters are only "weakly identified", and cannot be estimated using the Ethiopian experimental data.[42]

Fechner error specification: I now consider the second error model used by Harrison and Rutström (2008) – the Fechner error model – which was originally formulated by Fechner (1966) and popularized by Hey and Orme (1994).[43] This error specification was designed to analyze decisions between pairs of lotteries, as present in experimental decision problems using the MPL design; however, I apply this error specification to the data from the Ethiopian experiment by coding choices in a binary format, as described later.[44]

In the Fechner error model, first the expected utility (EU) for each potential choice – the right (R) or left (L) lottery – in a lottery pair p of decision problem d is calculated (as a function of r) using the CRRA utility function described in Equation (1) and the expected utility formulation in Equation (2) (see Appendix B for a detailed derivation of this error specification). Then, the latent index

[42] For all the maximum likelihood estimations of this chapter in which I encountered non-convergence of the optimization procedure, I also used scans (in MATLAB) to shed light on the cause of non-convergence. First, I calculated the log-likelihood function for a range of different parameter values, and found that this function was flat over a wide range of parameter values. Further, I also conducted maximum likelihood estimations using a range of different initial (starting) parameter values. I found that many different initial values produced markedly different maximum likelihood parameter estimates; additionally, these different estimates corresponded to the same (maximized) value of the log-likelihood function. In other words, the log-likelihood function was flat over a range of maximum likelihood parameter estimates – multiple parameter estimates produced the same (maximized) value of the log-likelihood function, and there is insufficient information to identify unique maximum likelihood estimates. This was confirmed by graphs of the log-likelihood functions. In addition, this was the case for a number of different optimization algorithms (for example, sequential quadratic programming, Newton-Raphson and Berndt-Hall-Hall-Hausman). Thus, the analysis indicates that the cause of the non-convergence is the nature of the experimental data at hand, rather than a failure of the numerical optimization procedure or non-identification of the model itself.

[43] Wilcox (2008) argues that the Fechner error specification should be favoured over the Luce error specification in the analysis of experimental choice.

[44] It is important to note that there a number of other stochastic error specifications – which place the error at different points in the decision-making process – that have been used in the experimental literature (for a review, see Wilcox 2008). However, the "white noise" error approach considered here, which generates the Luce and Fechner models and has its foundations in the psychophysical literature, has been widely used in similar studies for estimating and evaluating alternative models of choice under risk; further, it is likely to represent the most appropriate method to model errors in this context (Hey and Orme 1994, Loomes et al. 2002).

∇EU^d_{pFech} for each lottery pair is calculated as follows:

$$\nabla EU^d_{pFech} = \left(\frac{EU^d_{pR} - EU^d_{pL}}{\mu} \right) \tag{16}$$

where EU^d_{pR} and EU^d_{pL} represent the expected utilities for the right and left lotteries, respectively, in lottery pair p of decision problem d. The latent index is then linked to observed choices using the standard normal cumulative distribution function Φ (Harrison and Rutström 2008). Using this "probit" link function, the probability of choosing lottery R from a pair of lotteries p is given by:

$$\text{Prob(choose lottery R)} = \Phi(\nabla EU^d_{pFech}) \tag{17}$$

and thus:

$$\text{Prob(choose lottery L)} = 1 - \Phi(\nabla EU^d_{pFech}) = \Phi(-\nabla EU^d_{pFech}) \tag{18}$$

These probabilities can then be used to construct the log-likelihood function. Once again, as $\mu \rightarrow 0$, the specification collapses to the deterministic EUT model, where there are no stochastic errors in decision-making and the choice is strictly determined by the expected utility of the two lotteries, with the option providing greater utility always chosen; also, larger values of μ imply greater errors in the decision-making process and greater deviation of observed choices from true preferences, or a greater extent of deviation in choice probabilities from the difference between the expected utilities of two lotteries in a binary question.

However, since this error model is designed to analyze binary lotteries, I cannot apply it directly to the data from the experiment, since in each decision problem there are six, rather than two, choice options. Therefore, to apply this specification to the experimental data at hand, I follow the strategy used by Dave et al. (2010) in a similar context – I first convert the six lotteries in each decision problem into five hypothetical lottery pairs that result from comparing adjacent lotteries; each choice made by subjects is then transformed into a set of five binary decisions made for these hypothetical lottery pairs. For the insurance problems, this translates to:

Decision 1: Gamble B v. Gamble A
Decision 2: Gamble C v. Gamble B
Decision 3: Gamble D v. Gamble C
Decision 4: Gamble E v. Gamble D
Decision 5: Gamble F v. Gamble E

where A, B, C, D, E and F are the six choice options in the Ordered Lottery Selection decision problems. The decisions are arranged so that the option (within a particular pair) which involves the purchase of more insurance is located on the left, and Decision 5 considers the two options which involve the most insurance purchase. For the two indemnity insurance problems, this implies that the riskier lottery is always located on the right, and Decision 5 corresponds to the highest level of risk aversion.[45] To mirror this ordering, the benchmark decision problem is transformed as follows:

Decision 1: Gamble E v. Gamble F
Decision 2: Gamble D v. Gamble E
Decision 3: Gamble C v. Gamble D
Decision 4: Gamble B v. Gamble C
Decision 5: Gamble A v. Gamble B

A binary decision is assigned a value of 1 if the right lottery would be chosen and 0 if the left lottery would be chosen, based on information from the observed choices of subjects in the Ordered Lottery Selection problems of the Ethiopian experiment. For example, if a subject chose Gamble D in an insurance

[45] The binary decisions in the MPL problem of Holt and Laury (2002) are ordered in a similar manner, with switching from the left lottery to the right lottery at a later (lower) decision in the table indicative of greater risk aversion.

problem, this implies a coding of 0, 0, 0, 1, 1 for binary decisions 1 through 5, respectively. Similarly a choice of Gamble D in the benchmark decision problem implies a coding of 0, 0, 1, 1, 1. This transformation of the data makes the implicit assumption that if participants were given such a list of binary decisions in a MPL decision problem, they would not make any inconsistent (or intransitive) choices; that is, they would only switch at one point from choosing the left lottery to choosing the right lottery, and not repeatedly switch back and forth (or switch from the right lottery to the left lottery). Thus, the choices in the hypothetical binary lotteries are inferred from the actual observed choices in the experimental decision problems by imposing monotonicity on revealed preferences and enforcing transitivity (Andersen et al. 2006b, Charness et al. 2013). These are also the assumptions made by variants of the MPL design which impose a single switch point, that is, which ask participants to only choose the point at which they would switch from the left to the right lottery.[46]

After transforming the choices into binary decisions using this procedure, I construct the latent index for each lottery pair (as described in Equation (16)), and compute the corresponding probabilities of choosing the right and left lotteries (as specified in Equations (17) and (18)). I then use these probabilities to construct the log-likelihood function for the choices in the benchmark problem:

$$\ln L^{Fech}(r, \mu; \boldsymbol{y}) = \sum_{a=1}^{378} \sum_{p=1}^{5} [(\ln \Phi(\nabla EU^{B}_{pFech}) \times \mathbf{I}(y^{B}_{ap} = R)) + (\ln \Phi(-\nabla EU^{B}_{pFech}) \times \mathbf{I}(y^{B}_{ap} = L))] \tag{19}$$

[46] Such variations of the MPL procedure have been used, for example, by Tanaka et al. (2010) and Liu (2008).

Similarly, the log-likelihood for the insurance decision problems is:

$$
\begin{aligned}
&\ln L^{Fech}(r,\mu;\boldsymbol{y}) = \\
&\sum_{a=1}^{136}\sum_{p=1}^{5}[(\ln\Phi(\nabla EU_{pFech}^{T_{IM}}) \times \mathbf{I}(y_{ap}^{T_{IM}} = R)) + (\ln\Phi(-\nabla EU_{pFech}^{T_{IM}}) \times \mathbf{I}(y_{ap}^{T_{IM}} = L))] \\
&+\sum_{a=1}^{258}\sum_{p=1}^{5}[(\ln\Phi(\nabla EU_{pFech}^{T_{IX}}) \times \mathbf{I}(y_{ap}^{T_{IX}} = R)) + (\ln\Phi(-\nabla EU_{pFech}^{T_{IX}}) \times \mathbf{I}(y_{ap}^{T_{IX}} = L))] \\
&+\sum_{a=1}^{120}\sum_{p=1}^{5}[(\ln\Phi(\nabla EU_{pFech}^{T_{GM}}) \times \mathbf{I}(y_{ap}^{T_{GM}} = R)) + (\ln\Phi(-\nabla EU_{pFech}^{T_{GM}}) \times \mathbf{I}(y_{ap}^{T_{GM}} = L))] \\
&+\sum_{a=1}^{242}\sum_{p=1}^{5}[(\ln\Phi(\nabla EU_{pFech}^{T_{GX}}) \times \mathbf{I}(y_{ap}^{T_{GX}} = R)) + (\ln\Phi(-\nabla EU_{pFech}^{T_{GX}}) \times \mathbf{I}(y_{ap}^{T_{GX}} = L))]
\end{aligned}
\tag{20}
$$

where $\mathbf{I}(\cdot)$ is the indicator function and $y_{ap}^{d} = R(L)$ denotes the implied choice of the right (left) lottery by individual a in the hypothetical binary lottery p of decision problem d; $\boldsymbol{y}$ denotes the vector of (binary) choices incorporated in the estimation. These log-likelihood functions can then be maximized with respect to r and μ to obtain the maximum likelihood estimates of these two parameters. By imposing the restriction of transitive choice and converting the data into binary form, I am able to estimate the error parameter μ, in addition to r. This highlights the possibility that the converted data provides more power for the accurate estimation of the error parameter than the original data, even though risk preferences are still classified into only six categories (based on the switch point in the binary lotteries implied by the actual observed choice). The results of these maximum likelihood estimations are presented in Table 9. Since each choice in the original data is converted into five binary decisions in the transformed data, there are 1890 decisions for the benchmark problem and 3780 decisions for the four insurance problems in the transformed data.

Table 9 shows that the estimated coefficient of relative risk aversion r is 1.254 for the benchmark problem and 0.877 for the insurance problems – further, these estimates are statistically significantly different from 0, once again indicating that participants in the experiment are risk averse. Additionally, the estimate of the

Table 9. *Maximum likelihood estimates of CRRA utility function with Fechner error*

Coefficient	Estimate	Standard Error	p-value	Lower 95% Confidence Interval	Upper 95% Confidence Interval
A. Benchmark Decision Problem					
r	1.254	0.030	0.000	1.196	1.312
μ	0.024	0.003	0.000	0.018	0.029
B. Insurance Decision Problems					
r	0.877	0.031	0.000	0.816	0.938
μ	1.453	0.336	0.000	0.795	2.111

Wald test p-values and 95% confidence intervals reported.
The log-likelihoods corresponding to the estimations in Panels A and B are -982.4 and -2569.3.

Fechner error parameter μ (0.024) for the decisions in the benchmark problem is statistically significantly different from 0 at the 1% level, implying a rejection of the hypothesis ($\mu = 0$) that there are no stochastic errors in the decision-making process and that observed choices are determined solely by true preferences (in this case, EUT with CRRA utility). However, the small magnitude of the estimate (relative to the no error value 0) implies that if EUT with CRRA utility is assumed to be the true decision-making process, there are relatively few errors made by participants in the benchmark problem. Crosetto and Filippin (2013) and Dave et al. (2010) also estimate a similar Fechner error model combined with the EUT-CRRA specification to assess the degree of noise in experimental decision-making. Following their interpretation, the relatively small noise parameter indicates that participants' decisions in the benchmark problem conform closely to the predictions of EUT, with a low degree of randomness in choice (that is, with relatively few choices attributed to errors in decision-making).

The estimate of μ for the insurance problems is quite large in magnitude (relative to the no error value of 0), and is statistically significant. This implies that if EUT with CRRA utility is assumed to represent the true decision-making process of subjects, there are a significant number of choices that cannot be explained by the difference in expected utilities between the right and left lotteries, indicating substantial stochastic errors (randomness) and considerable deviation of observed choices from true preferences; in other words, choice probabilities do not depend highly on the difference in expected utilities and subjects' choices in the insurance problems may not conform closely to the predictions of EUT-CRRA.

Further, the hypothesis of $\mu = 1$ is rejected for the benchmark decision problem (but not for the insurance problems). Thus, the results in Table 9 highlight the importance of accounting for and estimating the magnitude of stochastic variation in the decision-making process – not accounting for the noise parameter could lead to markedly different parameter estimates and thus contrasting descriptions of behaviour under risk (a point also noted by Loomes et al. 2002). That is, when μ is significantly different from 1, as is the case in the Fechner error estimation for the benchmark problem, the structural estimations without the μ parameter – which implicitly assume that $\mu = 1$ – may yield different, and possibly incorrect, estimates of the preference parameters. This is because these estimations provide the maximum likelihood estimates of the preference parameters (in the case of EUT-CRRA, r) given that the noise parameter is fixed at 1; however, if μ is significantly different from 1, the preference parameter estimates may be very different in the specifications involving the free noise parameter. For example, the results for the EUT-CRRA specification show that the estimated r for the benchmark problem is 0.537 when it is assumed that $\mu = 1$, but is 1.254 (indicating substantially greater risk aversion) when μ is considered to be a free parameter. As a result, the structural specifications involving the free noise parameter – that is, the Luce and Fechner error models described above – should be preferred to the less-flexible specification that assumes a fixed extent of errors in decision-making (that is, $\mu = 1$). The error specifications allow the magnitude of noise in decision-making to vary and be estimated, thus providing greater flexibility and a better indication of true risk preferences.

Additionally, it is important to note that the wide range of estimates of r (in the EUT-CRRA model) obtained in Tables 7, 8 and 9 is likely due to the different assumptions – especially those related to stochastic errors – made by the different specifications. The interval regression specification (Table 8) does not incorporate stochastic errors in decision-making, the structural specification in Table 7 incorporates stochastic errors but assumes a fixed extent of errors in decision-making (that is, μ=1), while the structural specification in Table 9 also incorporates stochastic errors in decision-making but allows the extent of these errors to vary and be estimated.

Once again, in Table 9, using Wald and likelihood ratio tests, the hypothesis that the coefficients of relative risk aversion in the benchmark and insurance decision problems are the same is rejected at the 1% significance level (as is the case for μ) – this provides further indication that different decision models may be generating choices in the benchmark and insurance problems (possibly due to differences in framing or complexity), and thus choices in these problems should be considered separately in estimations.[47] Bearing these factors in mind, I explore possible deviations from EUT-CRRA in the following sections.

5.3 Expo-power utility

In this section, I still assume EUT, but no longer assume CRRA utility. Instead I use the more flexible two-parameter expo-power (EP) utility function described in Equation (3). Using this utility function, the expected utility of each lottery is calculated (as a function of the two parameters) using objective probabilities, as described in Equation (2). The latent index ∇EU_i^d for each potential choice i in decision problem d is then constructed (as described in Equations (10) and (11)), and used to create the log-likelihood function (as specified in Equations (12) and (13)) – this is similar to the methodology used earlier for the EUT-CRRA functional form. The log-likelihood function can then be maximized with

[47] I obtain similar results for the Fechner estimations – for both the benchmark and insurance decision problems – when I use a logistic, rather than normal, cumulative distribution function to link choice probabilities to the latent index (as described in Equations (17) and (18)). This is the case for both the CRRA and EP decision models. As noted by Train (2009), when a logistic function is used with the latent index specified in Equation (16), the model collapses to the Luce error model for the two-option case.

Table 10. *Maximum likelihood estimates of expo-power utility function*

Coefficient	Estimate	Standard Error	p-value	Lower 95% Confidence Interval	Upper 95% Confidence Interval
A. Benchmark Decision Problem					
r_{EP}	0.655	1.856	0.726	-2.983	4.293
α_{EP}	-0.368	5.946	0.952	-12.022	11.286
B. Insurance Decision Problems					
r_{EP}	1.324	0.417	0.001	0.507	2.141
α_{EP}	0.177	0.037	0.000	0.104	0.250

Wald test p-values and 95% confidence intervals reported.
The log-likelihoods corresponding to the estimations in Panels A and B are -628.9 and -1344.3.

respect to r_{EP} and α_{EP} to obtain maximum likelihood estimates of these two EP parameters. Once again, I separately analyze decisions in the benchmark problem and the insurance problems, and the results of these maximum likelihood estimations are listed in Table 10.

For the benchmark decision problem, I obtain an estimate of r_{EP} of 0.655, though this estimate is statistically insignificantly different from 0, indicating that I cannot reject the hypothesis of CARA ($r_{EP} = 0$). The estimate of α_{EP} for the benchmark problem is negative – however, this estimate is also statistically insignificantly different from 0, implying that I cannot reject the hypothesis of CRRA ($\alpha_{EP} = 0$) for decisions in the benchmark problem. The standard errors for both parameter estimates are relatively large – one possible reason for this is the fact that there are fewer observations (only half) in the benchmark problem than in the insurance problems, and this lower sample size increases the likelihood of obtaining estimates that are not statistically significant.[48]

For the insurance decision problems, the estimate of r_{EP} is 1.324 and is statistically significantly different from 0 – thus, the hypothesis of CARA ($r_{EP} =$

[48] As a result of these large standard errors, I am unable to reject the hypothesis (at the 10% significance level) that each of the parameters is the same for the benchmark and insurance decision problems, using Wald tests.

0) is rejected at the 1% significance level in favour of decreasing absolute risk aversion ($r_{EP} > 0$). The estimate of α_{EP} is 0.177 and is also statistically significantly different from 0, implying a rejection of the hypothesis of CRRA ($\alpha_{EP} = 0$) in favour of increasing relative risk aversion ($\alpha_{EP} > 0$). This indicates that there is evidence of increasing relative risk aversion over the domain of prizes in the insurance decision problems.[49] Holt and Laury (2002) and Lee (2008) also obtain positive estimates of both EP parameters in their analyses of data from experiments involving university students, implying that the EP utility function exhibits increasing relative risk aversion and decreasing absolute risk aversion. Additionally, a rejection of the hypothesis of CRRA suggests that the CRRA functional form used in Section 5.1 may not accurately describe the choices made by participants in the insurance decision problems.

I also conduct estimations of the EUT-EP model combined with the Fechner error parameter, using the transformed binary data (as described in the previous section).[50] The results of these estimations for the data from the benchmark and insurance problems are presented in Table 11.

As before, for the benchmark decision problem, the estimate of α_{EP} is statistically insignificantly different from 0, indicating that the hypothesis of CRRA ($\alpha_{EP} = 0$) cannot be rejected for decisions in the benchmark problem. However, in this case, the coefficient of r_{EP} is negative and statistically significant at the 1% level, implying a rejection of CARA ($r_{EP} = 0$) in favour of increasing absolute risk aversion ($r_{EP} < 0$) (Saha 1993).[51]

[49] It is important to remember that the lottery prizes in the experiment represent significant amounts of money, as average earnings in the experiment (40 birr) is two to three times the daily income from casual farm labour in the region. Thus, the domain of prizes represents an important one over which inhabitants, especially farmers (who make up a majority of the experiment sample), regularly make decisions involving risk.

[50] Once again, I encountered non-convergence of the optimization procedure when I attempted to estimate the EUT-EP model in combination with the Luce error parameter.

[51] Increasing absolute risk aversion has little conceptual justification for our sample. However, it is important to note that in these specifications, the estimates of the preference parameters represent the parameter values conditional on EUT-EP plus error representing the true data-generating process. Thus, if subjects' true preferences diverge substantially from EUT-EP and the noise parameter estimates imply substantial randomness in choice, the preference parameter estimates may not provide an accurate reflection of true risk attitudes.

Table 11. *Maximum likelihood estimates of expo-power utility function with Fechner error*

Coefficient	Estimate	Standard Error	p-value	Lower 95% Confidence Interval	Upper 95% Confidence Interval
A. Benchmark Decision Problem					
r_{EP}	-0.964	0.365	0.008	-1.680	-0.247
α_{EP}	0.001	0.002	0.454	-0.002	0.004
μ	39.12	60.64	0.519	-79.73	158.0
B. Insurance Decision Problems					
r_{EP}	-0.288	0.025	0.000	-0.337	-0.240
α_{EP}	0.031	0.002	0.000	0.027	0.036
μ	0.538	0.054	0.000	0.432	0.643

Wald test p-values and 95% confidence intervals reported.
The log-likelihoods corresponding to the estimations in Panels A and B are -981.3 and -2186.9.

However, for the insurance decision problems, α_{EP} is positive and statistically significant at the 1% level, implying a rejection of the hypothesis of CRRA ($\alpha_{EP} = 0$) in favour of increasing relative risk aversion ($\alpha_{EP} > 0$). As is the case for the benchmark problems, r_{EP} is negative and statistically significant at the 1% level, implying a rejection of the hypothesis of CARA ($r_{EP} = 0$) in favour of increasing absolute risk aversion ($r_{EP} < 0$). Harrison and Rutström (2008) also obtain positive estimates of α_{EP} and negative estimates of r_{EP} in some of their expo-power estimations involving experimental data from the United States. The sign and magnitude of the estimates differ significantly between this model, which incorporates the free error parameter, and the previous model (which assumes $\mu = 1$).[52] This highlights the importance of taking into account the magnitude of errors in experimental decision-making, as this may have important implications for the inferred behaviour under risk of subjects.

Additionally, for the insurance decision problems, the noise parameter μ is estimated to be around 0.5 – it is statistically significantly different from 0 at the 1% level, implying a rejection of the hypothesis of no errors in decision-

[52] Further, the hypothesis of $\mu = 1$ is rejected for the insurance decision problems (but not for the benchmark problem).

making ($\mu = 0$). However, for the decisions in the benchmark problem, the error parameter is imprecisely estimated (that is, with a relatively large standard error and wide confidence interval) and is statistically insignificantly different from 0.

In the next section, I relax some of the assumptions underlying EUT and explore the importance of loss aversion in generating the choices in the insurance problems.

5.4 Loss aversion

In this section, I no longer assume that EUT describes the true decision-making process under risk of subjects in the experiment. I extend the specification to include the possibility of loss aversion. To do this, I consider a two-part power utility function – as described in Equation (4) – which is defined separately over gains and losses.[53] For now, I assume that there is no subjective probability weighting; however, I relax this assumption in the next section.

Harrison and Rutström (2008) note that a key operational weakness of this two-part power utility function (and of prospect theory) is the need to specify what the reference point is, and that this reference point is subjective and contextual in nature. For example, Abeler et al. (2011) and Andersen et al. (2006a), analyzing experimental data from Germany and the United States, respectively, find that the reference point is based on expectations of experimental earnings – in line with the model of Koszegi and Rabin (2006, 2007) – and this expectation-based reference point influences effort provision by individuals. Further, the effect of this specification ambiguity is magnified by the fact that individuals' preferences, and thus estimates of the model parameters, depend crucially on the reference point, as this determines what outcomes are evaluated as losses and gains (Munro

[53] While the original formulation of prospect theory allowed for different coefficients of curvature for gains and losses, I use the same α_P for the functions defined over gains and losses – this is common in empirical investigations of the two-part power utility function and prospect theory (Harrison and Rutström 2008). Further, numerous researchers (for example, Tversky and Kahneman 1992, Harrison and Rutström 2009), obtain estimates of the coefficients of curvature that are statistically indistinguishable for gains and losses.

and Sugden 2003). Indeed, Sprenger (2010) notes that reference-dependent preferences with different reference points provide notably different accounts of decision-making behaviour, particularly in relation to financial and insurance purchase decisions. Additionally, numerous non-experimental studies have found that the reference point in such reference-dependent models has an impact on crucial real-world decisions, such as the decision regarding the amount of labour to supply (for example, Crawford and Meng 2011)[54], as well as consumption and financial trading decisions (Munro and Sugden 2003). Kahneman and Tversky (1979) also recognize the importance and subjectivity of the reference point in their original presentation of prospect theory, and note that utility should be treated as a function of two arguments, the reference point and the magnitude of change from that reference point.

Most experimental studies assume an exogenous reference point – generally chosen on the basis of the frame presented by the lottery outcomes (prizes) – in similar reference-dependent decision models (for example, Liu 2008, Tanaka et al. 2010). However, Munro and Sugden (2003) and Falk and Knell (2004) note that there is substantial evidence against assuming an exogenously-given reference point, and highlight the importance of considering the reference point to be endogenous, as it is likely to be actively chosen by individuals.

Further, Andersen et al. (2006a) and Harrison and Rutström (2008) stress the importance of taking a structural perspective when estimating such reference-dependent models and allowing the data to determine what the reference point is – given the importance of the reference point in characterizing and understanding risk attitudes, assuming the wrong reference point could yield misleading preference parameter estimates and incorrect inferences regarding behaviour under risk. Indeed, Andersen et al. (2006a) find that even though researchers commonly use the experimenter-induced lottery frame to select exogenous reference points, there are often "homegrown" reference points that

[54] Crawford and Meng (2011), analyzing the labour supply decisions of New York City cabdrivers, estimate reference points as sample proxies for earnings expectations.

subjects bring to the task, which are not related to the framing of the lottery but are based on prior experiences or beliefs about the experimental task. Therefore, I define utility to be a function of the reference point χ as well as the lottery prize x; I follow the strategy of Andersen et al. (2006a) and Harrison and Rutström (2008) and consider the reference point to be endogenous, and attempt to estimate it jointly with the other parameters in the model (α_P and λ). Though this increases the number of parameters to be estimated, in this manner, I allow the data to determine the reference point used by subjects in the decision problems, in order to be able to make consistent statements about behaviour under risk.[55]

Using objective probabilities to weight the utilities from different outcomes, the latent index and log-likelihood function are constructed in the same manner as done for the EUT-CRRA specification (described in Section 5.1), except that the two-part power utility function is used to evaluate lottery prizes instead of the CRRA utility function. However, when I attempted to maximize the log-likelihood function for choices in the benchmark problem, the optimization routine did not converge (though it did converge for choices in the insurance problems). This can be attributed to the fact that I am now attempting to estimate three, rather than two parameters (as before), and there are insufficient decisions and insufficient variation in the size of the lottery prizes to permit the joint estimation of all three parameters.

It is important to note that the decision problems of the Ethiopian experiment were not designed for the sole purpose of estimating and testing between different decision theories; rather, they were designed to shed light on insurance purchase decisions. This is in contrast to other experiments which contain decision problems specifically designed – with choice options, lottery prize values and corresponding probabilities selected – to identify and accurately elicit the parameter values of non-EUT decision models. For example, Tanaka et al.

[55] Ideally, given the subjective nature of the reference point, I would allow for heterogeneity in the reference point across individuals – for example, by assuming a different reference point for each individual. However, for parsimony and estimation purposes, I assume a homogenous reference point and thus aim to estimate the reference point for the average subject in the sample.

(2010) conduct an experiment in rural Vietnam which is explicitly designed to elicit the three parameters of the CPT specification directly from participants' choices in three MPL decision problems – that is, the decision problems are designed so that, with strict functional form assumptions, any combination of responses in the three problems determines a particular interval for each of the three prospect theory parameters.[56] Similarly, the experimental design created by Barr (2007) includes MPL decision problems in the gain and loss domain that are specifically designed to test the presence of probability weighting, while the experiment conducted by Humphrey and Verschoor (2004a) in rural Uganda involved decision problems that were designed to test between expected utility theory, rank-dependent utility theory and prospect theory.

Thus, it is important to note that the decision problems of the Ethiopian experiment (analyzed in this chapter) are not likely to provide the same power for accurately estimating the parameters of the different decision models (and in particular, the error and non-EUT models that incorporate multiple parameters) as those comprising the above-mentioned experiments; this low power is a likely cause of the non-convergence of the optimization procedure when estimating the two-part power utility model, in addition to the lower precision associated with the Ordered Lottery Selection elicitation method (as noted earlier).[57]

Therefore, for the estimation of the two-part power utility specification using the Ethiopian data, I combine the choices from all five decision problems. However, Camerer (2005) points out that it may be inappropriate to view χ as a fixed parameter that does not vary with the context of the decision and framing of the decision problem. Additionally, Munro and Sugden (2003) note that individuals adjust their reference points quickly in response to changes in the status quo, endowment or framing. Therefore, since the benchmark problem and insurance problems were framed very differently, it is likely that the reference point used by participants varies between these problems – to account for this, I assume

[56] Liu (2008) implements the same experimental design in rural China.

[57] I follow Barr's (2007) usage of "power" in the context of experimental decision problems, and she makes a similar point regarding the relative power of different types of decision problems.

Table 12. *Maximum likelihood estimates of two-part power utility function*

Coefficient	Estimate	Standard Error	p-value	Lower 95% Confidence Interval	Upper 95% Confidence Interval
All Decision Problems					
χ_B	1.681	2.593	0.516	-3.401	6.763
χ_T	17.076	1.091	0.000	14.938	19.214
λ	3.029	0.147	0.000	2.741	3.317
α_P	0.511	0.055	0.000	0.403	0.619
$H_0 : \lambda = 1$			0.000		

Wald test p-values and 95% confidence intervals reported.
The log-likelihood corresponding to this estimation is -1931.0.

that there are different endogenous reference points for the benchmark decision problem and the insurance decision problems, and estimate both these reference points (jointly with the two other model parameters) using maximum likelihood.

As a result, I maximize the log-likelihood function with respect to four parameters – the reference point for the benchmark decision problem χ_B, the reference point for the insurance problems χ_T, the coefficient of loss aversion λ and the coefficient of curvature α_P. Thus, I assume that only the reference point differs for decisions in the benchmark and insurance problems, while the other two parameters are the same for all decision problems. This combined dataset provides more choices (1,134) and greater variation in lottery prizes, thus enabling the estimation of all four parameters using maximum likelihood. The results of this estimation are presented in Table 12.

The results indicate that the estimate of the reference point for B, χ_B, is 1.68 birr, and this estimate is statistically indistinguishable from 0. The lowest prize in the benchmark problem is 15 birr – therefore, the estimated reference point is below the entire distribution of payoffs in B.[58] In other words, relative to this reference point, there are only gains in B, and no losses, and only the utility function defined over gains is relevant for this decision problem. Further, as Andersen et al. (2006a) point out, when the reference point is 0, the two-part power utility specification collapses to one which assumes that utility is defined solely over final prize outcomes. In the benchmark decision problem, all outcomes are positive monetary prizes – thus, the two-part power utility function collapses to the one-part power utility function defined over gains, that is, $U(x) = x^{\alpha_P}$. This, however, is simply a form of CRRA utility, which has been used in numerous experimental studies (for example, Andersen et al. 2006a, Harrison and Rutström 2009).[59,60] Additionally, since I have assumed no subjective probability weighting, this specification is an EUT formulation – these results are in line with the hypothesis that EUT with CRRA utility is the decision-making process that generates choices in the benchmark decision problem, and that loss aversion plays no role in these decisions.

For the insurance decision problems, the reference point is estimated to be around 17 birr, and is statistically significantly different from 0.[61] Further, as expected, the reference points in the benchmark and insurance problems differ significantly – using Wald and likelihood ratio tests, the hypothesis that $\chi_B = \chi_T$ is rejected at the 1% significance level. The lottery prizes in the insurance problems range from 0 birr to 75 birr; thus, the estimated reference point for the insurance problems (χ_T) is at the lower end of the spectrum of lottery prizes.

[58] Note that it is possible, as observed by Crosetto and Filippin (2013) and Andersen et al. (2006a), that even for decision problems framed completely in the gain domain (like the benchmark problem) or the loss domain (like the insurance problems), a reference point can be endogenously determined by the outcomes in the task, and gains and losses evaluated against this reference point; thus, loss aversion could be triggered and exhibited in pure gain or loss domain decision problems. However, for a pure gain domain decision problem, it is possible that the likelihood may want to move the estimated reference point lower and lower, or higher and higher, for spurious reasons.

[59] The coefficient of relative risk aversion in this particular CRRA specification is α_P: $\alpha_P < 1$ corresponds to risk averse behaviour, $\alpha_P = 1$ to risk neutral and $\alpha_P > 1$ to risk loving. Thus, the estimate of 0.511 for α_P implies that participants exhibit risk aversion in the benchmark problem.

[60] Harrison et al. (2010) also note that when the part of the utility function defined over losses is dropped, the two-part power utility specification collapses to the CRRA specification.

[61] This is almost identical to the estimate obtained when using only the choices in the insurance decision problems for the estimation.

One interesting question to explore is whether this reference point is determined by the framing of the insurance decision problems or is "homegrown", that is, brought to the tasks by the subject and unrelated to the framing of the problems. As noted by Munro and Sugden (2003) and Abeler et al. (2011), it is crucial to study what determines the reference point – this facilitates an increase in the economic applicability of reference-dependent decision theories.[62] To analyze this issue, I test whether the estimated endogenous reference point is statistically distinguishable from three *a priori* candidate values of the reference point that are commonly considered to be induced by the framing of the problem: 0, which implies that participants determine the gain and loss frames directly from the sign of the prize (as suggested by Andersen et al. 2006a)[63]; 65 birr, the initial endowment provided in the insurance decision problems (as suggested by Kahneman and Tversky 1979, Harrison and Rutström 2008)[64]; and 25 birr, which is the amount obtained in the "safe" (or certain) choice option in the benchmark problem (option A) and in the indemnity insurance problems (option F), where 25 birr is obtained with certainty regardless of the token selected (as suggested by Abeler et al. 2011). Thus, in other words, participants can obtain 25 birr with certainty in these decision problems by choosing the "safe" (risk-free) choice option.[65] In an experiment designed to analyze the determinants of the reference point, Abeler et al. (2011) find that a similar "fixed" payment influences expectations of experimental earnings, which in turn determine the reference point. Further, Crosetto and Filippin (2013) note that the initial endowment in loss domain tasks and the amount in the certain (risk-free) choice option of Binswanger-style lotteries represent likely reference points for decision-making

[62] Further, Abeler et al. (2011) also note that if reference points are assumed on a case-by-case basis without any analysis of where they come from, "models of reference-dependent preferences might explain behavior not because of their structural assumptions but because of this additional degree of freedom".

[63] Liu (2008) and Tanaka et al. (2010) assume a reference point of 0 when using a similar two-part power utility function to analyze the risk attitudes of subjects in experimental decision problems.

[64] Munro and Sugden (2003) note that in an experimental setting, the initial endowment (65 birr, in this case) represents an individual's status quo position and thus is a likely candidate for his reference point.

[65] Note that in the index insurance problems, there is no such corresponding "safe" choice option, as none of the options in these decision problems provides full insurance against prospective losses.

that are determined by the framing of the problem.

The results in Table 12 indicate that the estimate of 17 birr is statistically significantly different from 0 at the 1% level. Using a Wald test, I find that this estimate is statistically distinguishable from 25 birr at the 1% significance level, implying a rejection of the hypothesis that $\chi_T = 25$; thus, there is evidence that the reference point is not equal to the amount that can be obtained with certainty (that is, the payoff in the "safe" choice option) in the benchmark and indemnity insurance decision problems. Similarly, the estimated reference point is statistically distinguishable from 65 birr at the 1% level. This provides evidence, in accordance with the results of Andersen et al. (2006a) and Harrison and Rutström (2008), that the reference point for the insurance problems is "homegrown" rather than determined by the framing of these problems.[66]

Additionally, the estimate of λ is approximately 3. Using a Wald test, I can conclude that this estimate is statistically distinguishable from 1 at the 1% level. Since the hypothesis of $\lambda = 1$ (that is, no loss aversion) is rejected in favour of $\lambda > 1$, there is evidence of loss aversion exhibited in the insurance decision problems. Thus, there is evidence that the decisions in these problems are not appropriately characterized by EUT, which does not incorporate loss aversion. The magnitude of the λ estimate is similar to that obtained by Tanaka et al. (2010) (2.63) and Liu (2008) (3.47) using experimental data from rural Vietnam and China, respectively. Additionally, Humphrey and Verschoor (2004a) also find evidence of loss aversion among experimental subjects in East Uganda, while Barr and Packard (2005) – utilizing a non-parametric analysis – find that participants in their experiment conducted in Peru exhibit loss aversion (rather than pure risk aversion).

The estimate obtained here, however, is larger than that obtained by many prospect theory analysts – primarily using data from developed countries –

[66] Further, Crosetto and Filippin (2013) note that even if the reference point is "homegrown", it could be affected by the outcomes in the task – in line with this observation, the evidence indicates that the reference points used by subjects in the benchmark and insurance problems differ significantly.

who find that $\lambda \approx 2$ (Harrison and Rutström 2008). This relatively large estimate of λ implies a considerable degree of loss aversion, which could drive individual preferences and have important consequences for decision-making, particularly in relation to production and investment strategies (Fafchamps 2009). For example, Liu (2008), using a similar two-part power utility function, finds that higher loss aversion leads to the delayed adoption of a new technology (in this case, genetically modified high-yield varieties of cotton crops) by a sample of Chinese farmers, and thus lowers wealth accumulation from technological innovations.[67]

Further, the estimate of α_P, the coefficient which determines curvature of the utility function, is 0.511; this estimate is statistically different from both 0 and 1 at the 1% level. Since the estimated $\alpha_P < 1$, the estimate indicates that the utility function is (mildly) concave over gains and (mildly) convex over losses; that is, participants are risk averse over gains and risk seeking over losses relative to the reference point of 17 birr. This estimate of α_P is similar to the corresponding estimate of 0.48 obtained by Liu (2008) and 0.6 obtained by Tanaka et al. (2010), using experimental data from low-income regions. However, it is slightly lower than the estimate of around 0.7 obtained by Harrison and Rutström (2008) and Harrison and Rutström (2009) in their analyses of experimental data from the United States.

The estimates of χ_T, λ and α_P imply that, as illustrated in Figure 1 and as predicted by Tversky and Kahneman (1992), the utility function has an "S shape" and is steeper for losses than for gains near the reference point of 17 birr.[68] These estimates can also explain the finding of increasing relative risk aversion (for the insurance problems) obtained using the EP utility function in Section 5.3 – a kink at the reference point of 17 birr (which is a consequence of significant loss

[67] In addition, Liu (2008) also finds that higher values of the curvature parameter (α) lead to slower adoption of the new technology.

[68] Harrison and Rutström (2008) emphasize that caution must be exercised in extending these results to other samples or datasets, as they depend closely on the reference point, which differs significantly between contexts.

aversion) combined with a utility function that is concave over gains and convex over losses could generate the finding of increasing relative risk aversion over the domain around the kink.[69]

The results in this section suggest that for the benchmark problem, EUT with CRRA utility may be the appropriate decision model to describe choices. However, for the insurance problems, I find evidence that utility is defined over deviations from a non-zero reference point rather than over final lottery prizes, and that subjects evaluate gains and losses differently (exhibiting loss aversion) – this implies a rejection of EUT-CRRA as the latent decision-making process generating choices in the insurance problems. I find instead that a two-part power utility function with loss aversion, or an expo-power utility function with increasing relative risk aversion (as estimated in Section 5.3), better describe the decisions in the insurance problems. In order to shed light on the suitability of each of these two competing decision models – which are non-nested – for describing the decisions of subjects in the insurance problems, I use a mixture model approach, as done by Harrison and Rutström (2009), Harrison et al. (2010) and Conte et al. (2011).

5.5 Mixture Model

The mixture model approach does not assume that a single model or decision theory generates all the choices. Rather, it recognizes the possibility that several latent behavioural processes may co-exist in a population and indeed that heterogenous theories may co-exist in the same sample; as a result, it allows decisions to be generated by more than one decision model and lets the data determine the extent of support for each model (Harrison and Rutström 2009,

[69] The two-part power utility model combined with the Luce and Fechner error specifications could not be estimated due to non-convergence of the optimization procedure, likely due to the relatively large number of parameters to be estimated. Dave et al. (2010) also encounter convergence problems when attempting to estimate the error models using Ordered Lottery Selection data that is transformed into binary choice data (as described in the Section 5.2), particularly for expo-power and power utility specifications; additionally, Crosetto and Filippin (2013) encounter non-convergence even when attempting to estimate such models using MPL data.

Conte et al. 2011).[70]

Following this mixture model approach, I specify a grand likelihood function in which both decision models – expo-power (EP) utility defined over final monetary prizes and two-part power utility (PU) defined separately over gains and losses – exist and have different (non-zero) weights. Using maximum likelihood, I can then estimate the weight assigned to each theory in generating choices in the insurance decision problems – this weight parameter is interpreted as indicating the fraction of choices in the sample better described by the corresponding decision model. Therefore, while homogeneity within a given theory (that is, all participants have the same preference parameter values, as implied by the representative agent assumption) is still assumed, I now allow heterogenous theories (preference functionals) to generate the choices of individuals. The grand log-likelihood function is[71]:

$$\begin{aligned} \ln L^{MM}(r_{EP}, \alpha_{EP}, \chi_T, \lambda, \alpha_P, \pi^{EP}; \boldsymbol{y}) &= \sum_{y_c=1}^{756} \ln[\pi^{EP} L^{EP}_{y_c}(r_{EP}, \alpha_{EP}; y_c) \\ &\quad +(1-\pi^{EP}) L^{P}_{y_c}(\lambda, \chi_T, \alpha_P; y_c)] \end{aligned} \tag{21}$$

where $\boldsymbol{y}$ is the vector of the 756 choices in the insurance problems, and $L^{EP}_{y_c}$ and $L^{P}_{y_c}$ are the conditional likelihood functions (corresponding to observed choice y_c) of the expo-power and power utility models, respectively, which are calculated as described in Sections 5.3 and 5.4. π^{EP} denotes the probability that the EP model is correct, that is, the probability that an observation is better characterized

[70] Harrison and Rutström (2009) observe that ignoring this possibility could lead to incorrect conclusions about the applicability of each theory, while Conte et al. (2011) find that using a mixture model adds substantially to the explanatory power of their estimates. Therefore, they argue that the mixture model could provide a better metric for comparing the suitability of different non-nested models for describing the data, as compared to formal likelihood ratio tests of one model against the other (as used by Harless and Camerer 1994). Such tests, including the Vuong (1989) test and Clarke (2003) test, implicitly assume a single data-generating process.

[71] Following Harrison and Rutström (2009) and Harrison et al. (2010), I do not use mixture model specifications in which participants are characterized entirely by one model or the other, in order to allow for the possibility that subjects behave in accordance with the EP model for some choices and the PU model for other choices, which is possible even with relatively homogenous decision problems. However, with the pooling of subjects and estimation of the average preference parameter values, this is expected to make little difference to the estimates.

by the EP model (as compared to the PU model); thus, $\pi^P = 1 - \pi^{EP}$ denotes the probability that the PU model is correct. Thus, these weighting parameters reflect the proportion of choices in the data better described (or accounted for) by each model and provide an indication of how well each model fits the data.

Harrison and Rutström (2009) and Harrison et al. (2010) obtain maximum likelihood estimates of all the utility function parameters and the weighting parameter (in this case, π^{EP}), by maximizing the grand log-likelihood function jointly with respect to these parameters. However, when I attempted to maximize the above grand log-likelihood function with respect to all the parameters, the optimization routine did not converge. Harrison and Rutström (2009) also point out that one difficulty associated with mixture models is the joint estimation of the weighting and the model parameters – if each model has some chance of explaining the data, then mixture models are characterized numerically by relatively flat likelihood functions.[72]

Therefore, I fix the values of the utility function parameters to the maximum likelihood estimates obtained in Sections 5.3 and 5.4 for data from the insurance problems, and maximize the grand log-likelihood function with respect to only π^{EP}.[73,74] Using this estimation strategy and the choices in the insurance problems, I obtain an estimate of π^{EP} of 0.092, implying that π^P is estimated to

[72] Harrison et al. (2010), using a much larger sample of 4,248 experimental choices, find that their sample size is not large enough to enable the estimation of the mixture model with a substantial set of covariates for the risk preference parameters; this further highlights the convergence issues associated with maximum likelihood estimations of the mixture model.

[73] The parameter values of the utility functions used in the estimation of the mixture model are: $r_{EP} = 1.324$, $\alpha_{EP} = 0.177$, $\chi_T = 17.076$, $\lambda = 3.029$, $\alpha_P = 0.511$. These values correspond to an EP function with increasing relative risk aversion and a PU model with loss aversion and an S-shaped value function.

[74] In accordance with the strategy used by Harrison and Rutström (2009), the maximum likelihood estimation actually provides estimates of the log odds in favor of one model over the other. They note that, denoting the log odds as κ, one can recover the probability of the EP model being correct as $\pi^{EP} = 1/(1 + \exp(\kappa))$. Thus, this non-linear function of κ can be easily calculated from the estimates, and the "delta method" can be used to provide estimates of the standard errors and p-values (Oehlert 1992). Since $\pi^{EP} = \frac{1}{2}$ when $\kappa = 0$, the standard p-value of the estimate of κ provides the estimate for the null hypothesis $H_0 : \pi^{EP} = \pi^P = \frac{1}{2}$.

be 0.908.[75] These results suggest that in the sample of decisions in the insurance problems, only 9.2% of the observations are better characterized by expo-power utility with increasing relative risk aversion, while 90.8% of the choices are accounted for by the two-part power utility function with loss aversion, assuming that the data can only be generated by these two latent behavioural processes.

Further, the hypothesis that $\pi^{EP} = \pi^{P} = \frac{1}{2}$ (that is, each of the two models has equal support in the data) is rejected at the 1% level, using Wald and likelihood ratio tests. The estimates of both probabilities, π^{EP} and π^{P}, are also significantly different from 0. Thus, while there is support for each model in the data, the high estimated probability π^{P} provides some evidence that most decisions in the insurance problems can be appropriately described by the two-part power utility function with loss aversion, rather than by expo-power utility with increasing relative risk aversion. Thus, in this case, while there is still some support for the EP model, assuming that the PU model generates all the decisions in the insurance problems may be an appropriate approximation. Since the two-part power utility function represents a deviation from EUT, the results in this section provide further evidence for the rejection of EUT as the behavioural model generating the choices in the insurance problems.

I also conducted Vuong (1989) and Clarke (2003) tests to compare the EP and PU models. In both cases, the EP specification is convincingly rejected (at the 1% significance level) in favour of the PU specification, implying that the PU model is closer to the true data-generating process (and thus fits the data better) than the EP model.[76]

[75] The estimate of the log odds ratio κ is 2.291, while the standard error is 0.105. Additionally, the log-likelihood corresponding to this estimation is -1295.0.

[76] It is important to note that the Vuong (1989) and Clarke (2003) tests, as well as the mixture model estimations, are tests of specific parametric forms of the decision theories, rather than general tests of the theories themselves. Therefore, these tests enable me to compare the suitability of different parametric forms for describing the experimental data.

5.6 Rank-dependent utility and cumulative prospect theory

In this section, I consider the possibility of subjective probability weighting, which is another important deviation from EUT. I assume that the choices of subjects are generated by rank-dependent utility (plus a fixed magnitude of noise), which allows for the weighting of cumulative probabilities. As described in Section 2.3, I consider two single-parameter functional forms for the probability weighting function ω – the Tversky-Kahneman and Prelec functions – which convert objective cumulative probabilities into subjective decision weights.

Note that using data only from the benchmark decision problem B, the probability weighting parameter (that is, γ in the Tversky-Kahneman function and ϕ in the Prelec function) is not identified. This is because identification of the probability weighting parameter requires variation in the probabilities associated with the different outcomes in the decision problem. However, there is no such variation in the benchmark decision problem, as all outcomes in this decision problem are associated with the same probability (0.5), implying the same cumulative probability weights for a given value of the weighting parameter. Indeed, Harrison and Rutström (2008) note that it is impossible to make inferences about probability weighting using versions of the Ordered Lottery Selection elicitation procedure that restrict probabilities to 0.5.[77]

In order to identify and accurately estimate the probability weighting parameters, variation is required in the objective probabilities associated with different lottery prizes – therefore, I only estimate the RDU specifications using data from the insurance decision problems (756 choices). I combine each of the two weighting functions with the three utility functions considered earlier – CRRA, EP and PU. After computing the utility of each prize outcome using these functions, the rank-dependent utility RDU_i^d from lottery i in decision problem d is calculated (as a function of the preference and weighting parameters) using the decision weights,

[77] Since the original experimental lotteries of Binswanger (1980) also involve only probabilities of 0.5, they do not enable the identification of the probability weighting parameter.

where the decision weights are computed as described in Equations (6) and (7). As illustrated for EUT in Section 5.1, the log-likelihood function $\ln L^{RDU}$ is constructed as follows:

$$rdu_i^d = \exp(RDU_i^d) \tag{22}$$

$$\nabla RDU_i^d = \frac{rdu_i^d}{\sum_{j=1}^{6} rdu_j^d} \tag{23}$$

$$\begin{aligned}\ln L^{RDU}(\boldsymbol{r}, \gamma; \boldsymbol{y}) \quad = \quad & \sum_{a=1}^{136} \ln(\nabla RDU_{y_a^{T_{IM}}}^{T_{IM}}) + \sum_{a=1}^{258} \ln(\nabla RDU_{y_a^{T_{IX}}}^{T_{IX}}) \\ & + \sum_{a=1}^{120} \ln(\nabla RDU_{y_a^{T_{GM}}}^{T_{GM}}) + \sum_{a=1}^{242} \ln(\nabla RDU_{y_a^{T_{GX}}}^{T_{GX}})\end{aligned} \tag{24}$$

where y_a^d is the observed choice of individual a in decision problem d and $\boldsymbol{y}$ denotes the vector of observed choices incorporated in the estimation; $\boldsymbol{r}$ is the vector of parameters in the utility function being analyzed and γ is the probability weighting parameter. The log-likelihood function $\ln L^{RDU}$ can then be maximized with respect to $\boldsymbol{r}$ and γ jointly in the case of Tversky-Kahneman probability weighting, and with respect to $\boldsymbol{r}$ and ϕ jointly in the case of Prelec probability weighting, to yield maximum likelihood estimates of these parameters. As noted earlier, when I combine cumulative probability weighting with the two-part power utility function, this provides a CPT specification.[78]

Originally, I intended to estimate all three utility functions – CRRA, EP and PU – in combination with each of the two probability weighting functions. However, I encountered non-convergence of the optimization procedure in the RDU estimation of the EP utility function, for both the Tversky-Kahneman and Prelec probability weighting functions. Therefore, I am only able to obtain

[78] While the original CPT specification of Tversky and Kahneman (1992) assumed a different coefficient of curvature and a different probability weighting parameter (within the same probability weighting function) for gains and losses, I assume that both these parameters are the same for gains and losses, as is common in empirical analyses of prospect theory (Harrison and Rutström 2008); further, as noted by Harrison and Rutström (2009), these differences are not critical to prospect theory.

maximum likelihood estimates of the RDU weighting model combined with the other two utility functions (CRRA and PU), and these estimates are reported in Table 13. In the estimations involving the two-part power utility function, I assume the reference point to be 17.076 birr, which is the maximum likelihood estimate obtained in Section 5.4, and maximize the log-likelihood function jointly with respect to λ, α_P and the probability weighting parameter.[79]

For the CRRA function, the estimates of the coefficient of relative risk aversion (r) obtained are around 1 for both the Tversky-Kahneman and Prelec RDU specifications, and are statistically significantly different from 0. These are slightly higher than the estimate of 0.774 obtained under EUT (that is, without probability weighting), but are in line with moderate risk aversion. Additionally, the estimate of γ is 2.156 for the Tversky-Kahneman probability weighting function and the estimate of ϕ is 2.821 for the Prelec probability weighting function. Parameter estimates greater than 1 imply S-shaped probability weighting, with convex weighting for low probabilities and concave weighting for high probabilities (see Figure 6). Furthermore, the S shapes are statistically significant – using Wald tests, the null hypothesis that there is no probability weighting ($\phi = 1$,$\gamma = 1$) is rejected at the 1% level for both probability weighting functions, and there is convincing evidence that $\phi > 1$ and $\gamma > 1$.

Since the hypothesis of no probability weighting is rejected, EUT is also rejected in favour of RDU – with an S-shaped probability weighting function – combined with CRRA utility. Though Gonzalez and Wu (1999) and Andreoni and Sprenger (2010) note that the probability weighting function is commonly found to be inverse S-shaped, especially for experiments conducted with standard samples in developed countries, the results obtained here are consistent with those of Humphrey and Verschoor (2004a), Humphrey and Verschoor (2004b) and

[79] This is because when I attempted to estimate the reference point, jointly with the other parameters, the optimization routine did not converge, given the relatively large number of parameters to be estimated. On the other hand, the results in Chapter 3 indicate that all the RDU models (including the RDU-EP model) as well as all the associated preference parameters (including the reference point) are accurately estimated using MPL data from an experiment conducted in Brazil.

Table 13. *Maximum likelihood estimates of rank-dependent utility models*

Coefficient	Estimate	Standard Error	p-value	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Insurance Decision Problems					
A. CRRA Utility with Tversky-Kahneman Probability Weighting Function (RDU-CRRA)					
r	0.938	0.030	0.000	0.880	0.995
γ	2.156	0.318	0.000	1.532	2.779
$H_0 : \gamma = 1$			0.000		
B. CRRA Utility with Prelec Probability Weighting Function (RDU-CRRA)					
r	0.968	0.060	0.000	0.850	1.086
ϕ	2.821	0.451	0.000	1.938	3.704
$H_0 : \phi = 1$			0.000		
C. Two-part Power Utility with Tversky-Kahneman Probability Weighting Function (CPT)					
(Reference Point = 17.076 birr)					
λ	2.831	0.421	0.000	2.006	3.655
α_P	0.424	0.036	0.000	0.353	0.495
γ	1.072	0.100	0.000	0.875	1.268
$H_0 : \lambda = 1$			0.000		
$H_0 : \gamma = 1$			0.475		
D. Two-part Power Utility with Prelec Probability Weighting Function (CPT)					
(Reference Point = 17.076 birr)					
λ	2.706	0.404	0.000	1.915	3.497
α_P	0.446	0.039	0.000	0.370	0.522
ϕ	1.133	0.091	0.000	0.955	1.311
$H_0 : \lambda = 1$			0.000		
$H_0 : \phi = 1$			0.142		

Wald test p-values and 95% confidence intervals reported.
The log-likelihoods corresponding to the estimations in Panels A, B, C and D are -1378.5, -1350.5, -1290.5 and -1289.7.

Harrison et al. (2010), who find evidence of S-shaped probability weighting in laboratory experiments (framed in the abstract) conducted with non-standard samples of subjects in developing countries. These results lend credence to the possibility, acknowledged by Humphrey and Verschoor (2004a), that there may be some systematic feature of relatively uneducated individuals in developing countries that reverses the shape of the probability weighting function.[80] Indeed, Banerjee (2000) notes that the poor are subject to different pressures from the rest of the population, and thus may have a distinct decision-making process from the non-poor, leading to very different choices under risk. Harbaugh et al. (2002) also find evidence of S-shaped probability weighting in an experiment involving children, another class of subjects with low levels of formal education.

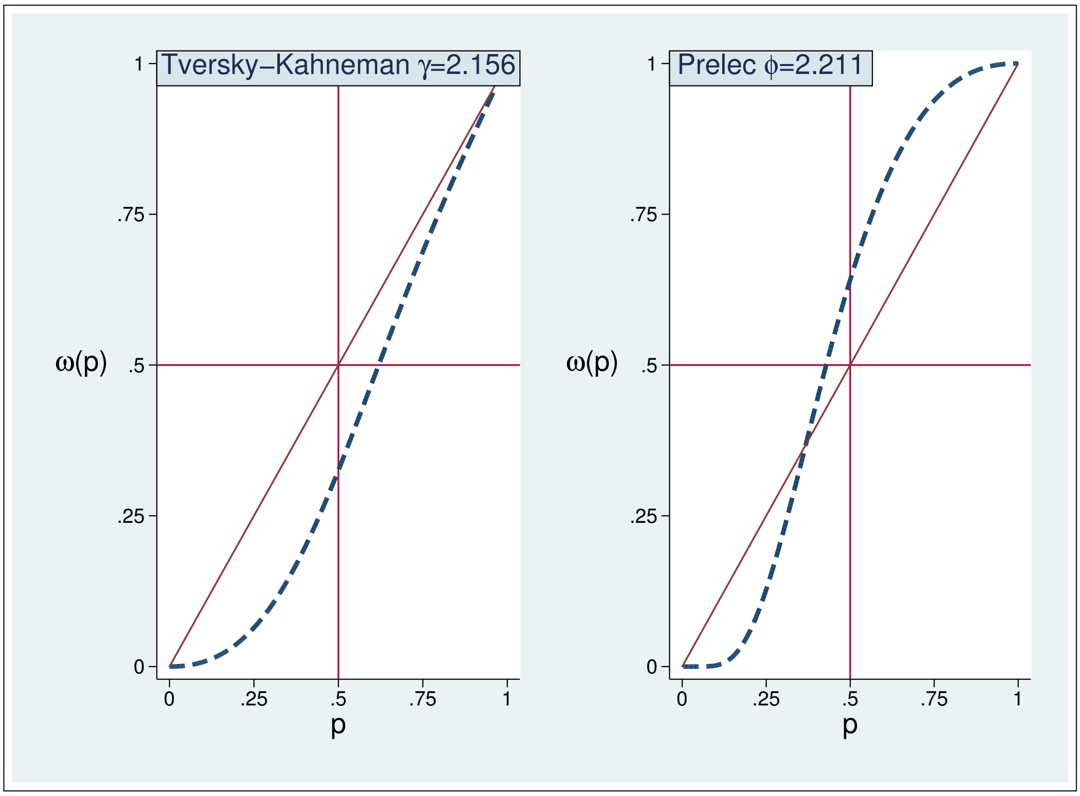
In Figure 6, which graphically displays the shapes of the probability weighting functions implied by the weighting parameter estimates, the presence of considerable underweighting is observed for virtually the entire range of probabilities when using the Tversky-Kahneman weighting function – this phenomenon is also observed by Harrison et al. (2010) in their analysis of experimental data, and is noted by Gonzalez and Wu (1999) to be a characteristic of the Tversky-Kahneman weighting function (when $\gamma > 1$).[81,82] For the Prelec probability weighting function, the estimates imply that $w(p) = p$ at a

[80] In poor rural economies where a significant fraction of inhabitants live close to the subsistence level, informal insurance arrangements exist where individuals are provided assistance by others (for example, family members or neighbours) when they experience low probability adverse shocks (Fafchamps 2003). Conversely, during low probability favourable events, these individuals are required to give a substantial fraction of their income away to insure others. While they cope individually with ordinary (high probability) events, their outcomes are significantly attenuated during extraordinary (low probability) events; in extraordinary circumstances, poor inhabitants of rural regions may feel that they are not responsible for their actions (which is not likely to be the case for developed country inhabitants). As a result, it is possible that individuals in rural regions place lower weights on low probability events in their decision-making, and follow these decision-making rules or heuristics in laboratory experiments as well. This is a possible cause of the S-shaped probability weighting – that is, underweighting of low probabilities and overweighting of high probabilities – observed in developing country lab experiments.

[81] Harrison et al. (2010) postulate that this significant underweighting of a wide range of probabilities may reflect the pessimism of subjects – possibly due to the difficult economic conditions in rural Ethiopia – which cause them to behave as if outcomes in the problems have much less chance of occurring than they actually do.

[82] Harrison et al. (2010) note that an important limitation of the Tversky-Kahneman weighting function is that it does not allow independent specification of location and curvature. That is, it has a fixed point, at $1-1/e = 0.63$ for $\gamma > 1$, below which the function is convex and above which it is concave.

Figure 6. *Estimated probability weighting functions for the RDU-CRRA specification*



probability of approximately 0.37 – there is an underweighting of probabilities ($w(p) < p$) below this point and an overweighting of probabilities ($w(p) > p$) above it (the weighting function is also convex below this point and concave above it). Additionally, for both probability weighting functions, there is diminishing marginal sensitivity to changes from probabilities near the center of the distribution, making the functions steeper in the middle and flatter towards the ends (Humphrey and Verschoor 2004b).

For the cumulative prospect theory specifications (that is, probability weighting combined with the two-part power utility function), the estimates of λ are 2.831 when I consider the Tversky-Kahneman probability weighting function and 2.706 when I consider the Prelec weighting function. Both these estimates are statistically different from 1 at the 1% level, implying that there is significant loss aversion – these estimates are quite similar to the estimate of around 3 obtained in Section 5.4 when I assumed no probability weighting, and are not in line with EUT. Similarly, the estimates of the curvature parameter α_P obtained (around

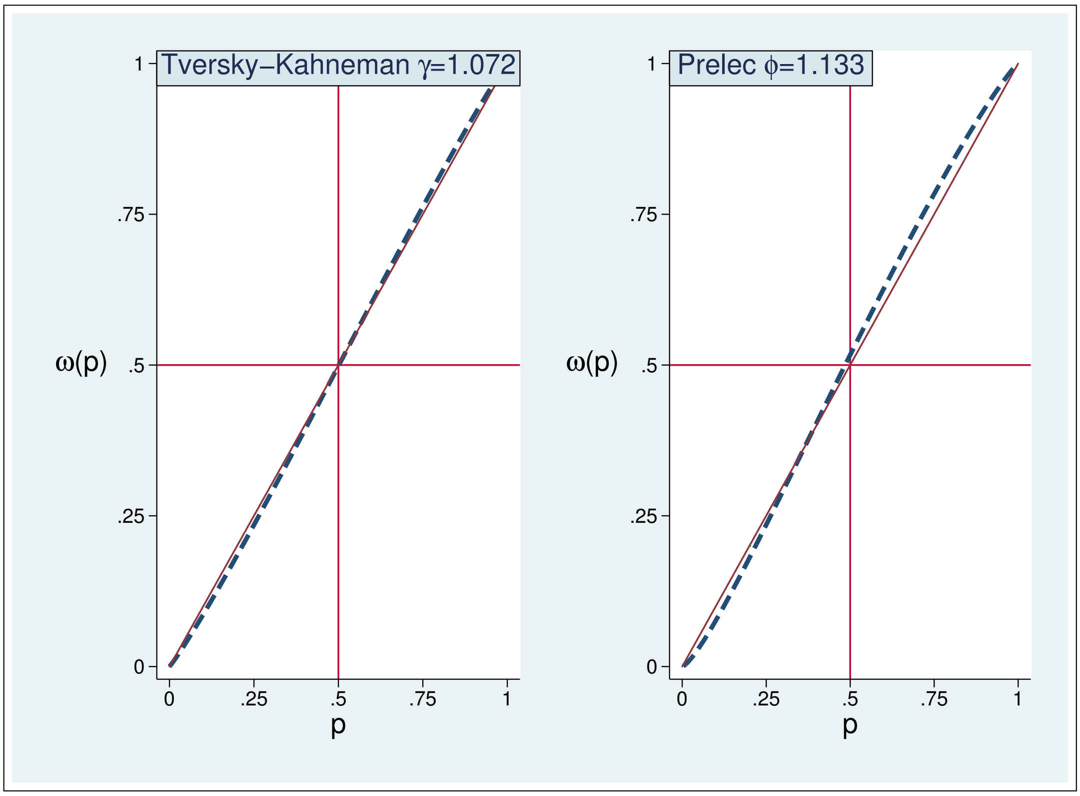
0.45 in both cases) are also similar to those obtained in Section 5.4.

For the CPT specifications, the estimates of the probability weighting parameter are very close to 1. For both the Tversky-Kahneman and Prelec weighting functions, Wald tests indicate that the formal hypothesis of no probability weighting ($\gamma = 1$ and $\phi = 1$) cannot be rejected at the 10% significance level. Thus the estimates, along with Figure 7, provide evidence of no, or very minimal, probability weighting, unlike in the case of the RDU-CRRA specification. This provides an indication that the CPT specification collapses to the two-part power utility specification with loss aversion but no probability weighting, for describing choices in the insurance problems. Therefore, it seems that the data features of the decisions in the insurance problems that were being captured by the probability weighting parameter in the RDU-CRRA specification are now being encompassed by the two-part power utility function with loss aversion but no probability weighting (as considered in Section 5.4).[83] Further, these results once again suggest that EUT does not appropriately describe subjects' choices in the insurance decision problems, with RDU-CRRA and two-part power utility (with loss aversion but little or no probability weighting) more likely to represent the true decision-making process generating choices in these problems.

In order to test the suitability of each of these two non-nested models – RDU with CRRA utility and two-part power utility – for describing the decisions in the insurance problems, I utilize the traditional Vuong (1989) and Clarke (2003) tests. These tests have been extensively used in the experimental economics literature to compare the fit of non-nested decision models (for example, Loomes et al. 2002, Blavatskyy and Pogrebna 2010, Wilcox 2010). Using these tests and data from the insurance problems, the RDU-CRRA specification is rejected in favour of the two-part power utility specification (with no probability weighting) at the 1% significance level, implying that the two-part power utility specification is closer to the true data-generating process generating choices in the insurance problems;

[83] Along similar lines, Harrison et al. (2010), comparing RDU-CRRA and EUT-CRRA specifications, find that probability weighting substitutes, to some extent, for the concavity of the utility function when explaining the data.

Figure 7. *Estimated probability weighting functions for the RDU-PU specification*



in other words, the two-part power utility specification better describes the true decision-making process of individuals in the insurance problems and thus fits the data better.[84]

In the next two sections I analyze two noteworthy patterns observed in the experimental dataset in the context of the empirical analysis of decision-making under risk – namely, the high take-up of index insurance in decision problems T_{IX} and T_{GX}, and the significant difference in choices between the numerically-identical decision problems B and T_{IM}.

[84] It would also be appropriate to use a mixture model, as described in Section 5.5, in this context. However, when I attempted to estimate this mixture model, the optimization routine did not converge, both when I allowed the decision model parameters and the mixture parameter (which indicates the fraction of choices in the data better described by each model) to vary, and when I fixed the decision model parameters to those values estimated in earlier sections, with the mixture parameter as the only free parameter.

6 Explaining the High Take-up of Index Insurance

According to the theoretical model of Clarke (2011), under expected utility theory and (risk averse) constant relative risk aversion, it is never optimal to pay an insurance premium of more than 3 birr in the individual index insurance problem T_{IX} and more than 12 birr in the group index insurance problem T_{GX} (see Tables 2 and 4).[85,86] In other words, no expected utility maximizing decision-maker with preferences satisfying risk aversion and constant relative risk aversion would ever pay an insurance premium of more than 3 birr in T_{IX} and more than 12 birr in T_{GX} (for any value of the coefficient of relative risk aversion $r > 0$) – this is because if the individual is averse enough to risk to want to purchase the cover, he would also be averse enough to the downside basis risk (that is, the risk that the net transfer from insurer to policyholder does not match the incurred loss) to limit the size of the hedge.[87,88] It is important to note that such a limit on the amount of index insurance purchase is a feature of all expected utility preferences satisfying risk aversion and decreasing absolute risk aversion (DARA), and not just constant relative risk aversion (and by extension, not just

[85] Note that Clarke's (2011) theory only applies to risk averse individuals. However, in the context of poor inhabitants of rural Ethiopia and decisions involving significant amounts of money (mean realized earnings from the Ethiopian experiment was two to three days' income from casual farm labour), risk loving behaviour has little conceptual justification. Additionally, the results in this chapter (as well as those in Chapter 2) indicate that subjects in the Ethiopian experimental sample exhibit moderate risk aversion in the decision problems. Indeed, Charness and Viceisza (2012) note that risk loving behaviour has rarely been observed in experimental studies, and when it is observed, is indicative of a lack of understanding of the decision problems on the part of subjects (leading to "frivolous" responses), rather than reflective of true risk preferences. Thus, while this theory applies only to risk averse individuals, it is highly likely that the entire sample – comprising of experiment participants who are relatively poor – falls within this category.

[86] In this section, I refer to the wider class of utility functions that satisfy constant relative risk aversion (within the framework of expected utility theory) as expected utility theory constant relative risk aversion preferences. Thus, the theory and the results in this section apply to this wider class of functions (or preferences). This is in contrast to the preceding sections of this chapter, in which I referred to only a specific constant relative risk aversion functional form as EUT-CRRA utility.

[87] Theorem 4 of Clarke (2011) states that for risk averse decreasing absolute risk aversion preferences – to which class risk averse CRRA preferences belong – there is an upper bound to the optimal level of index insurance purchase.

[88] In the case of T_{IX}, for example, the downside basis risk is the probability that a blue bag (good index) is chosen – so there is no insurance payout – but a yellow (bad individual outcome) token is picked from the blue bag.

the particular EUT-CRRA functional form considered earlier in this chapter) (Clarke 2011). Further, Clarke (2011) considers the DARA upper bound to be a normatively sound basis for advice from an objective financial advisor and notes that, in the experiment, the DARA upper bound for index insurance purchase is only slightly less restrictive than the CRRA bound.[89] Thus, the theory of Clarke (2011) provides a one decision problem test for (risk averse) CRRA (and DARA) in experimental settings.

However, in the two index insurance decision problems, approximately 57% of the subjects purchased more insurance than would be optimal for any expected utility maximizer with (risk averse) constant relative risk aversion preferences.[90] Relatively high take-up rates of index insurance have also been observed in developing country experiments by Hill et al. (2009) (weather indexed insurance) and Lybbert et al. (2010) (area yield indexed insurance). In this section, I analyze this high take-up of index insurance, building on a similar analysis in Clarke and Kalani (2011).

At first glance this high take-up appears to contradict what has been found in early field studies involving index insurance purchase – Giné et al. (2008) and Cole et al. (2009), for example, find that voluntary purchase of index insurance products in rural India is very low (and is particularly low for the poorest), despite the prevalence of considerable risk in agricultural production. However, this is not an appropriate comparison – whilst observed demand in pilot projects in the field has been low, it has not been demonstrated to be "too low", or lower than optimal for a constant relative risk aversion (or decreasing absolute risk aversion) decision-maker. To be able to say anything meaningful about whether observed index insurance purchase is "too low", rather than just saying that it is "low", we must at the very least be able to argue that a well-informed financial advisor

[89] For example, Theorem 4 of Clarke (2011) implies that no expected utility maximizer with preferences satisfying risk aversion and DARA over aggregate wealth would ever pay an insurance premium of more than 6 birr in T_{IX}.

[90] In particular, in T_{IX}, 84% and 66% of the participants purchased more index insurance than the upper bound implied by CRRA and DARA preferences, respectively.

(in this case, considered to be an expected utility maximizer with risk averse CRRA or DARA preferences) would advise a higher level of purchase. This turns out to be surprisingly difficult to do, particularly for real index insurance products where the contractual index is not perfectly correlated with the loss, and neither researcher nor consumer has a precise objective estimate of the joint probability distribution of losses and index claim payments – in such a situation, it is not possible to compute the value of optimal demand for this financial advisor (Clarke 2011).[91]

Therefore, for understanding whether the level of demand for index insurance is too low or too high, field experiments suffer from a critical problem: rational (or optimal) demand, by which I mean the demand that would be recommended by a well-informed financial advisor, is highly sensitive to beliefs about the joint distribution of the losses and index claim payments, and an objective, unambiguous belief about this joint distribution is extremely difficult to obtain in the field (Clarke 2011).[92] By contrast, in a lab experiment, the losses and index can be generated by a known randomization device with objective joint probability distribution, and it is therefore possible to calculate the optimal demand for a (risk averse) constant relative risk aversion decision-maker, which in turn allows the researcher to make clear normative statements about the level of observed demand. This highlights the importance of lab experiments for making normative statements about the level of index insurance demand. Unlike recent field experiments and field studies[93], the Ethiopian laboratory experiment was

[91] The optimal demand can be described as a function of the parameters and the probabilities of different events occurring (for example, the policyholder experiencing a bad individual outcome but receiving no index insurance payment). However, without precise knowledge of the joint probability distribution of losses and index claim payments, the value of optimal demand cannot be calculated, and thus researchers cannot conclude whether observed demand is greater or less than this optimal value.

[92] Optimal demand from an expected utility maximizer is highly sensitive to both the expected claim payment and the claim payment distribution conditional on a large loss having occurred. Even if a researcher had, say, 20 years of matched data for both losses and index claim payments, the latter conditional distribution could not be objectively estimated with any degree of accuracy. In practice researchers are likely to have relevant matched data for significantly fewer than 20 years, rendering normative statements impossible.

[93] These include the studies of Giné et al. (2008) and Cole et al. (2009), who analyze the demand for index insurance using data from pilot projects in India, as well as the work of Hill et al. (2011), who investigate the same subject using survey data from Ethiopia.

designed with one of its specific aims being to allow normative statements to be made about the level of index insurance purchase. Therefore, it may well be that the low demand for products reported by Giné et al. (2008) and Cole et al. (2009), analyzing data from a pilot project conducted in India, is rational – thus it would be erroneous to label this demand as “too low”. Rather, the observed demand may actually be “about right” or even “too high” (as found in the Ethiopian experiment), rather than “too low” (as compared to rational or optimal demand) (Clarke 2011).[94]

The finding of considerable choices in the index insurance problems that are not optimal for expected utility theory and constant relative risk aversion preferences implies a direct rejection of this decision model as the latent behavioural process generating the choices in these problems (if there is no allowance for errors in decision-making). This observation is in line with the results in previous sections, where I also find evidence suggesting that EUT may not be appropriate for describing choices in the insurance problems.[95]

This high take-up of index insurance in decision problems T_{IX} and T_{GX} can, however, be explained by the finding of S-shaped probability weighting combined with the CRRA utility functional form (obtained in Section 5.6). For the index insurance decision problems, S-shaped probability weighting – which implies underweighting of low probabilities and overweighting of high probabilities (or relatively less underweighting of high probabilities in the case of the Tversky-Kahneman weighting function) – corresponds to underweighting the effect of downside basis risk (which is associated with low probability outcomes) and thus the over-purchasing of index insurance. In other words, since within the framework of RDU, probability weights translate into decision weights in the

[94] Cole et al. (2009) use historic weather data for the products they report on to estimate that for every 1 rupee of farmer premium, the expected claim payment to farmers is between 0.19 and 0.59 rupee. Combined with plausible estimates for the degree of basis risk, optimal demand from a (risk averse) constant relative risk aversion decision-maker is indeed close to zero.

[95] Evidence against EUT is consistent with that obtained in recent lab experiments framed in the abstract and conducted with subjects from developing countries (Humphrey and Verschoor 2004a, Humphrey and Verschoor 2004b, Liu 2008).

decision-making process of participants, S-shaped probability weighting implies that subjects are likely to underweight downside basis risk (which represents a major drawback of index insurance from the viewpoint of policyholders and reduces the desirability of the product)[96] in their decision-making process. Thus, one would expect observed demand for index insurance policies from participants in the experiment to be higher, not lower, than that which would be indicated by EUT-CRRA.[97]

Therefore, this provides further evidence that the decisions made by participants in the index insurance problems may be appropriately described by RDU with S-shaped probability weighting combined with CRRA utility. Humphrey and Verschoor (2004a) also find evidence that RDU with S-shaped probability weighting suitably describes the decision-making process of farmers in their experiment conducted in rural Uganda; additionally, they note the probability weights could affect the size of subsidies required to encourage microcredit take-up and also various aspects of microinsurance contracts.

While S-shaped probability weighting can explain the high demand for index insurance in the Ethiopian experiment, Liu (2008), analyzing probability weighting using the experimental choices of Chinese farmers, finds that individuals who underweight small probabilities are less likely to adopt new varieties of cotton crops (in particular, Bt cotton), leading to slower wealth accumulation. This result is surprising, given that (Liu 2008) notes that the genetically-modified Bt cotton produces higher yields and a naturally-occurring pesticide, and thus is associated with a relatively small probability of crop failure – as a result, we would expect farmers who underweight low probability adverse outcomes, such as crop failure, to be more, not less, likely to adopt Bt cotton.

[96] For example, Hill et al. (2011), analyzing survey data from Ethiopia, find that downside basis risk significantly reduces the demand for index insurance.

[97] On the other hand, Wakker et al. (1997), analyzing experimental data from a standard sample of developed country university students, find evidence of significant aversion to basis risk in a set of "probabilistic insurance" decision problems, which are similar to the index insurance purchase problems in the Ethiopian experiment but include a lower probability of basis risk, and interpret this as providing evidence for inverse S-shaped probability weighting (which involves the overweighting of low probability downside basis risk).

However, Liu (2008) also finds that adoption decisions are determined by the risk perceived by farmers rather than the real risk, which could explain the above-mentioned result; further, conducting a within-subject analysis, she finds that only a very small fraction of individuals exhibit S-shaped probability weighting.

7 EXPLAINING THE DIFFERENCE IN CHOICES BETWEEN B AND T_{IM}

The numerical values of the final outcomes (prizes) are identical in the benchmark decision problem B and the individual indemnity insurance problem T_{IM}, as are the probabilities associated with these outcomes. Therefore, if the choices in both decision problems are assumed to be generated by the same decision model (and there are no errors), one would expect participants to make identical choices in both decision problems. However, most participants made different choices in both problems – of the 136 subjects who were involved in both decision problems, only 19 of them made the same choice in both problems.[98] Further, it does not seem that the divergence of choices between the two decision problems is random – of the 117 (or 86%) participants who made different choices in the two problems, 79 (or 68%) chose "safer" options (that is, those with less variance in the final outcomes) in the insurance problem as compared to the benchmark problem, indicating a systematic difference in choices.[99]

While other experimental studies conducted using university members (for example, Camerer 1989, Hey and Orme 1994) and developing country subjects

[98] If a subject made choices in B and T_{IM} that correspond to the same outcome payoffs and probabilities, I refer to this subject as one who made the "same choice" in both the decision problems. The options in T_{IM} are in reverse order to the corresponding options in B – for example, option A in T_{IM} has the same outcomes and probabilities as option F in B; thus, if a subject chose these two options, he essentially made the same choice in both problems.

[99] Using a one-proportion z-test and an exact binomial probability test, the hypothesis that the proportion of switching subjects who chose "safer" options in the insurance problem is equal to 0.5 – which would be the case if subjects who switched their choice between the problems did so entirely at random – is rejected at the 1% significance level. Therefore, it is unlikely that the deviation in observed choices between the two problems is solely due to random errors; instead, it is likely to be systematic.

(for example, Humphrey and Verschoor 2004a) also find evidence of within-subject variability over repeated identical problems, they find it at a much lesser extent (around 30% of subjects report different choices in identical problems in these experiments). Therefore, if B and T_{IM} are viewed as "identical" problems, the high within-subject variability in observed choices between the two problems would imply extremely large, and somewhat improbable, errors in the decision-making process of participants. However, I am circumspect about considering these two problems as "identical", even though the numerical values of the final outcomes (as well as the corresponding probabilities) are the same – as mentioned in Section 4, the framing and complexity of the benchmark decision problem differed significantly from that of the indemnity insurance decision problem.

These important differences between the two problems are likely to lead to *preference reversals* on the part of subjects, which could cause the high and systematic within-subject difference in choices observed in the two problems – that is, this difference in choices could be due to a change in the true decision-making process, rather than random errors. Preference reversal would lead to a systematic change in decisions due to a difference in the framing or nature of otherwise identical tasks. It implies a violation of the assumption of procedure invariance which is implicit in most economic theories of choice and states that underlying preferences (and, in the absence of errors, choices in problems with numerically-identical outcomes and associated probabilities) do not vary with the procedure employed in the decision problem to elicit preferences (Humphrey and Verschoor 2004a). Indeed, Humphrey and Verschoor (2004a) note that preference reversal (and thus the violation of procedure invariance) is one of the most robust patterns of behaviour found in lab experiments conducted in developed countries over the last three decades; they also find evidence of this phenomenon in their non-standard experimental sample of Ugandan farmers. Much of the evidence on the subject indicates that elicited risk preferences are not stable across elicitation procedures, as well as across the degree of complexity, context, domain and framing of decision problems (Crosetto and Filippin 2013, Charness et al. 2013).

Thus, the evidence of large and systematic within-subject deviations in decisions between B and T_{IM} could be explained by preference reversals (rather than random errors), and further suggests that the decisions in B and those in T_{IM} (and by extension, possibly all insurance problems) are generated by different decision-making processes, which is in line with the findings in earlier sections. In other words, the two decision problems could invoke different decision-making processes, thus leading to significantly different preferences elicited as well as the decisions in the two problems being generated by different decision-making processes.

As noted by Munro and Sugden (2003), individuals' preferences vary systematically according to the perceived reference point, which in turn is likely to vary by the status quo as well as framing. Further, the results in Section 5.4 indicate that the estimated reference point differed significantly between the benchmark and insurance problems, in line with theoretical predictions and the results of other experimental studies. Thus, if the reference point is endogenous and allowed to vary between decision problems, the two-part power utility model (described in Section 5.4) could explain the systematic difference in choices between the two problems – since utility is a function of deviations from the reference point, which could vary depending on the problem, and gains and losses from this reference point are evaluated differently (due to loss aversion), choices could differ between B and T_{IM} because decisions in both problems are generated by the two-part power utility specification with different reference points.[100]

To explore the possibility that variation in the reference point, rather than random errors, is the cause of the systematic difference in choices observed, I consider the different possibilities of reference points for the two problems that could, in combination with two-part power utility and loss aversion, generate this observed

[100] Since the probabilities attached to the outcomes are the same in B and T_{IM}, probability weighting cannot explain the difference in choices between the problems. However, difference in the reference point could explain this phenomenon, and thus I only consider the two-part power utility model and ignore probability weighting (and thus CPT) in this section.

pattern of choices. In other words, I explore whether the observed difference in choices can be attributed to changes in true preferences (preference reversals) through changes in reference points, rather than simply to errors in the decision-making process of participants.

If I assume a two-part power utility function with the same reference point for both problems (for example, 0 or the certain amount 25 birr), this would imply (ignoring errors) that participants should make the same choices in both problems, since the probabilities and final prizes are the same and, with the same reference point, utility from each of the prizes is the same. Thus, two-part power utility with the same reference point in both problems cannot explain the substantial and systematic difference in choices (likely even if errors are permitted). Further, the pattern of choices cannot be explained by two-part power utility with a reference point of 25 birr in B and two-part power utility with a reference point of 65 birr (the initial endowment for the insurance problems) in T_{IM}, as in this case, the associated loss aversion would imply "riskier", rather than "safer", choices in T_{IM}. This is because with a reference point of 65 birr in the insurance problem, the entire distribution of payoffs is below the reference point, and there are only losses (no gains); since the two-part power utility function is convex over losses (if $\alpha_P < 1$), this implies risk seeking behaviour over losses, and would imply "riskier" choices (with higher variance in outcomes) in the insurance problem as compared to the benchmark problem, for which the reference point is 25 birr and there are both gains (over which behaviour is risk averse) and losses.

On the other hand, a reference point of 0 for the benchmark problem and 17 birr for the insurance problem, as estimated in Section 5.4, could explain the pattern of within-subject variability in choices observed.[101] For the insurance decision

[101] Strictly speaking, in Section 5.4 I estimate a reference point of 1.68 for the two-part power utility function using data from the benchmark problem. However, this estimate is statistically insignificantly different from 0; further, assuming the reference point to be 1.68 or 0 would provide the same qualitative conclusions in this section – this is because in either case, the reference point is below the entire distribution of prizes in the benchmark problem.

problem, loss aversion – which is found to be considerable in magnitude (see Section 5.4) – generates a kink at the reference point of 17 birr, and makes the utility function steeper for losses than for gains around this reference point. This kink, which is in the lower part of the distribution of payoffs in the insurance problem, combined with the fact the utility function is S-shaped (convex over losses and concave over gains), generates more risk averse behaviour (less risk taking) around the kink, thereby leading to "safer" choices in the insurance problem than in the benchmark problem, which has a reference point of 0 and thus only gains (over which utility is concave) relative to this reference point. Therefore, the results and model estimates (in particular, the reference point estimates) obtained in Section 5.4 can explain the systematic difference in choices between B and T_{IM}.

However, as noted in Section 5.4, when the two-part power utility has a reference point of 0 (as in the benchmark problem), it reduces to a constant relative risk aversion expected utility theory specification defined over the final prize outcomes. Thus, the systematic within-subject variability in choices observed is in line with preference reversal between the benchmark problem and the insurance problem, where the decisions in the benchmark problem are generated by a constant relative risk aversion expected utility theory specification and those in the insurance problem are generated by two-part power utility with loss aversion and a reference point of 17 birr. This once again suggests that there are two different decision-making processes that generate the choices in the benchmark decision problem and the insurance decision problems, and thus there is likely a change in the decision-making process between these two types of problems; in other words, the evidence highlights the possibility that different decision-making processes are invoked.[102] Humphrey and Verschoor (2004a) also find evidence that the decision-making process of participants in their experiment is task-dependent and sensitive to the framing and nature of

[102] However, it remains possible that the choices are generated by two different regimes of preferences, with the same decision-making process.

the task – they conclude that different types of tasks probably invoke different cognitive processes in decision-makers and thus there may not be a single choice theory which provides an overall description of the decision-making behaviour of subjects in all tasks. Further, Crosetto and Filippin (2013) find that the risk preferences elicited in their experiment conducted with German university students are highly task-specific; indeed, they note that a large number of studies on the subject have found that different risk attitudes are elicited (from the same population) using different elicitation procedures (or different framing of decision problems), with evidence suggesting that different decision-making processes are also invoked. In addition, Harrison et al. (2010) note that particularly in developing countries, a wider range of heuristics may be adopted to make decisions. This suggests that preference reversals, rather than random errors, could provide an appropriate explanation of the systematic difference in choices between decision problems B and T_{IM}, and by extension, possibly between B and all the insurance problems (given that all the insurance problems are framed similarly).[103]

This raises the following question: what characteristics of, and differences between, decision problems B and T_{IM} (and indeed, possibly between the benchmark problem and all the insurance problems) in the experiment cause subjects' to invoke different latent decision-making processes? One possible reason is the difference in framing of the two problems – while B is a simple Binswanger-style lottery framed in the gain domain, T_{IM} is framed in the loss domain with an initial endowment of 65 birr. Indeed, Harbaugh et al. (2002) note that there could be significant framing effects which cause subjects to evaluate and respond to decision problems in the gain and loss domains differently (using different decision-making processes), while Charness et al. (2013) note that elicited risk attitudes are highly domain-specific. As noted by Crosetto and Filippin (2013), tasks framed explicitly in the loss domain are more likely to

[103] Schmidt and Hey (2004), analyzing experimental data from the United Kingdom, find that choice errors play only a minor role in explaining observed preference reversals.

trigger loss aversion (and induce greater importance of the reference point) – this could explain the finding that a two-part power utility function with loss aversion better describes subjects' choices in the insurance decision problems, which are framed in the loss domain, while EUT-CRRA better describes choices in the gain domain benchmark decision problem.

Alternatively, the difference in the decision models used by participants to evaluate B and T_{IM} may be a result of the differing complexity of the two decision problems; while B involved a simple lottery, T_{IM} involved a compound one. Charness et al. (2013) note that complexity is a crucial aspect of decision problems, which determines the decision-making process invoked and the risk preferences elicited. Further, B involved no insurance aspect while T_{IM} was presented as a problem about insurance purchase, in which participants had to choose between different amounts of insurance (and different premium payments) – therefore, it is possible that this difference in abstract and contextual framing leads to different decision-making processes being invoked and systematically divergent choices in the two problems. For example, it is possible that subjects choose "safer" options – those which correspond to greater insurance purchase – in T_{IM} because they face considerable risk and volatile income streams in the field, and want to signal their need for insurance. If this is the case, insurance (including microinsurance) providers in this region (and possibly in other low-income regions) could use such experimental decision problems involving an insurance aspect to gauge the demand for insurance products, as the elicited preferences over outcomes in these problems may shed light on insurance purchase decisions in the field. In addition, insurance providers and policymakers should be wary of using risk attitudes elicited from more commonly-used decision problems framed in the abstract, as these problems may invoke different decision-making processes from the framed insurance problems, and thus choices in these problems may not reflect true preferences over, and decisions regarding, insurance purchase.

Thus, the results in this section, combined with those in the previous section, indicate the possibility that decisions in the indemnity insurance problem T_{IM} are better characterized by the two-part power utility specification with loss aversion (and a reference point of 17 birr) and those in the index insurance problems are better characterized by RDU-CRRA with S-shaped probability weighting. Thus, both decision models may play a major role in generating the choices in the insurance decision problems, and these models better describe the decisions in the insurance problems than the EUT models.[104] However, as noted in Section 5.6, using the traditional Vuong (1989) and Clarke (2003) tests for non-nested models and data from the insurance problems, the RDU-CRRA specification is rejected in favour of the two-part power utility specification (with no probability weighting) at the 1% significance level, implying that the two-part power utility specification is closer to the true data-generating process generating choices in the insurance problems (that is, the two-part power utility specification better fits the data than the RDU-CRRA specification).

8 CONCLUSION

Laboratory experiments constitute an important complement to fieldwork – particularly for assessing risk preferences – and are powerful tools for policy design and evaluation (Falk and Heckman 2009, Harrison et al. 2011). They provide an environment which offers the researcher a degree of control that is unfeasible in the field, and thus present the opportunity for directly testing the validity of different decision theories. In this chapter, I use different maximum likelihood estimation procedures to analyze data from five decision problems in a lab experiment conducted in rural Ethiopia, in order to assess risk preferences and compare a wide range of different theories of choice. I

[104] These results also highlight the importance of loss aversion and probability weighting in explaining behaviour under risk. Indeed, Tanaka et al. (2010) note that evidence from numerous studies (including theirs) indicates that non-EUT decision models which incorporate these features better fit both experimental and field data than EUT models (particularly those that include just a single preference parameter); further, they find that very few subjects in their experiment make choices consistent with EUT, finding instead evidence of loss aversion and probability weighting.

test both nested and non-nested decision models and find that expected utility theory with CRRA preferences provides a good description of choices in the benchmark Binswanger-style decision problem. On the other hand, I find that a rank-dependent utility specification (specifically, CRRA utility with S-shaped probability weighting) or a two-part power utility function with loss aversion – rather than EUT – is more likely to represent the latent decision-making process generating choices in the four insurance decision problems. This is in line with the results of Humphrey and Verschoor (2004a) and Tanaka et al. (2010), who find that expected utility theory fails to convincingly describe the experimental decisions made by inhabitants of Uganda and Vietnam, respectively, but rank-dependent and reference-dependent models perform better. Further, the results in this chapter highlight the importance of additional research on the explanatory power of reference-dependent decision models (which is also noted by Munro and Sugden 2003). Additionally, further research on what determines the reference point – in particular, whether it is "homegrown" or induced by the framing of problem – is essential, as the reference point is a crucial component of reference-dependent models.

The decision problems of the Ethiopian experiment analyzed in this chapter were designed to elicit attitudes toward, and gauge the demand for, indemnity and index insurance using experimental data, which has received little attention in the experimental literature. This distinct insurance aspect added an important dimension to the decision problems and yielded some noteworthy observations – I find that the level of insurance take-up in the two index insurance decision problems is higher than would be optimal for any expected utility maximizer with preferences satisfying risk aversion and constant relative risk aversion (and in fact, for any expected utility maximizer with preferences satisfying risk aversion and decreasing absolute risk aversion); additionally, I find that there is considerable within-subject variation in the choices in the benchmark problem and individual indemnity insurance problem, in spite of these problems having numerically-identical final outcomes and associated probabilities. Further, I find that both these phenomena can be explained by the non-EUT decision models estimated, namely RDU-CRRA with S-shaped probability weighting and two-part power utility with loss aversion. However, Vuong (1989) and Clarke (2003) tests for non-nested models indicate that the two-part power utility specification (with loss aversion but no probability weighting) is closer (of these two) to the true data-generating process for the insurance problems. The results indicate that there are significant framing effects which cause different latent decision-making processes to be invoked for the benchmark and insurance problems, and these preference reversals lead to systematically different choices in the problems.

However, certain decision models could not be estimated using the Ordered Lottery Selection data of the Ethiopian experiment, likely due to the lower power and coarser characterization of risk attitudes provided by this elicitation procedure (as compared to the MPL method). Further, the probability weighting parameter in RDU and CPT models is not identified using data from Binswanger-style lotteries – such as the benchmark decision problem in this experiment – that restrict the probability associated with each outcome to be the same. Studying the tradeoffs – primarily between precision and comprehension – for the relatively complex MPL method and the simpler Ordered Lottery Selection procedure is an important avenue for future research, particularly for lab experiments conducted in developing countries with poorly-educated subjects. Research on this topic would shed light on the crucial question of which experimental procedure is more appropriate for accurately eliciting the risk preferences of individuals in these regions.

Another important avenue for further research is to investigate the external validity of risk preferences elicited in experimental tasks, and to assess the extent to which these elicited preferences affect vital decisions in real life. The current literature on this issue is inconclusive – while some studies find that experimentally-elicited risk preferences do predict crucial decisions in real life, such as insurance purchase and technology adoption decisions (for example, Liu 2008), others find evidence of no relationship between the two (for example, Hill and Nobles 2011). For example, it would interesting to analyze whether the risk preferences elicited in the Ethiopian experiment have a bearing on the insurance purchase decisions of the participants in a pilot program offering real index insurance products, which will be conducted with ERHS households in the near future by the International Food Policy Research Institute (IFPRI). An analysis of this question could provide important insights that aid the design and delivery of insurance products in this region.

Evaluating the risk preferences – and characterizing the decision-making process under risk – of low-income populations is an important component of the experimental research agenda, and is crucial for both testing predictions of decision theories and for effective policy design and evaluation. The analysis of data from lab experiments conducted in developing countries provides valuable insights into the behaviour under risk of individuals in these regions; in addition, such lab experiments complement the theoretical work and field studies on decision-making under risk, enabling a richer characterization of risk attitudes and contributing significantly to this vital research program.

REFERENCES

Abeler, J., A. Falk, L. Götte, and D. Huffman, "Reference points and effort provision," *The American Economic Review*, 2011, *101* (2), 470–492.

Andersen, S., G.W. Harrison, and E.E. Rutström, "Choice behavior, asset integration and natural reference points," *Department of Economics Working Papers*, 2006.

____ , ____ , M.I. Lau, and E.E. Rutström, "Elicitation using multiple price list formats," *Experimental Economics*, 2006, *9* (4), 383–405.

Andreoni, J. and C. Sprenger, "Certain and Uncertain Utility: The Allais Paradox and Five Decision Theory Phenomena," *Levine's Working Paper Archive*, 2010.

Attanasio, Orazio, Abigail Barr, Juan Camilo Cardenas, Garance Genicot, and Costas Meghir, "Risk Pooling, Risk Preferences, and Social Networks," *American Economic Journal: Applied Economics*, 2012, *4* (2), 134–167.

Banerjee, A.V., "The Two Poverties," *Nordic Journal of Political Economy*, 2000, *26* (2), 129–141. Reprinted in Dercon, Stefan (ed), "Insurance Against Poverty", Oxford University Press, 2004.

Barr, A., "Risk Pooling, Commitment, and Information: An experimental test of two fundamental assumptions," *The Centre for the Study of African Economies Working Paper Series*, 2003, p. 187.

____ , "Trust and expected trustworthiness: experimental evidence from zimbabwean villages*," *The Economic Journal*, 2003, *113* (489), 614–630.

____ and G. Genicot, "Risk Sharing, Commitment, and Information: An Experimental Analysis," *Journal of the European Economic Association*, 2008, *6* (6), 1151–1185.

Barr, Abigail, "Attitudes to Risk in Ghana: Field Manual," 2007.

____ and Truman G Packard, "Seeking solutions to Vulnerability in Old Age: Preferences, constraints, and alternatives for coverage under PeruÕs pension system," *The Centre for the Study of African Economies Working Paper Series*, 2005, p. 239.

Binswanger, H.P., "Attitudes toward risk: Experimental measurement in rural India," *American Journal of Agricultural Economics*, 1980, *62* (3), 395–407.

____ , "Attitudes toward risk: Theoretical implications of an experiment in rural India," *Economic Journal*, 1981, *91*, 867–890.

Blavatskyy, P.R. and G. Pogrebna, "Models of stochastic choice and decision theories: why both are important for analyzing decisions," *Journal of Applied Econometrics*, 2010, *25* (6), 963–986.

Botelho, A., G.W. Harrison, L. Pinto, E.E. Rutström, and P. Veiga, "Discounting in developing countries: a pilot experiment in Timor-Leste," *NIMA Working Papers 31*, 2005.

Camerer, C., "Bounded rationality in individual decision making," *Experimental Economics*, 1998, *1* (2), 163–183.

——, "Three Cheers–Psychological, Theoretical, Empirical–for Loss Aversion," *Journal of Marketing Research*, 2005, pp. 129–133.

Camerer, C.F., "An experimental test of several generalized utility theories," *Journal of Risk and Uncertainty*, 1989, *2* (1), 61–104.

—— **and T.H. Ho**, "Violations of the betweenness axiom and nonlinearity in probability," *Journal of Risk and Uncertainty*, 1994, *8* (2), 167–196.

Cameron, A.C. and P.K. Trivedi, *Microeconometrics: methods and applications*, Cambridge Univ Pr, 2005.

Campenhout, B. Van, B. D'Exelle, and E. Lecoutere, "Measuring Attitudes towards Risk in Rural Tanzania," mimeo 2008.

Charness, G. and A. Viceisza, "Comprehension and risk elicitation in the field: Evidence from rural Senegal," *IFPRI Discussion Paper*, 2012.

Charness, Gary, Uri Gneezy, and Alex Imas, "Experiential methods: Eliciting risk preferences," *Journal of Economic Behavior & Organization*, 2013.

Clarke, D. and G. Kalani, "Microinsurance decisions: evidence from Ethiopia," Technical Report, Working Paper, Microinsurance Innovation Facility 2011.

Clarke, D.J., "A Theory of Rational Demand for Index Insurance," Department of Economics Discussion Paper Series 572, University of Oxford 2011.

Clarke, K.A., "Nonparametric model discrimination in international relations," *Journal of Conflict Resolution*, 2003, *47* (1), 72.

Cole, Shawn A., Xavier Giné, Jeremy B. Tobacman, Petia B. Topalova, Robert M. Townsend, and James I. Vickery, "Barriers to Household Risk Management: Evidence from India," Working Paper 09-116, Harvard Business School 2009.

Collier, P. and J.W. Gunning, "Explaining African economic performance," *Journal of economic Literature*, 1999, *37* (1), 64–111.

Conte, Anna, John D Hey, and Peter G Moffatt, "Mixture models of choice under risk," *Journal of Econometrics*, 2011, *162* (1), 79–88.

Crawford, V.P. and J. Meng, "New York City Cab Drivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational Expectations Targets for Hours and Income," *The American Economic Review*, 2011, *101* (5), 1912–1932.

Crosetto, Paolo and Antonio Filippin, "A Theoretical and Experimental Appraisal of Five Risk Elicitation Methods," Technical Report, Friedrich-Schiller-University Jena, Max Planck Institute of Economics 2013.

Dave, C., C.C. Eckel, C.A. Johnson, and C. Rojas, "Eliciting risk preferences: When is simple better?," *Journal of Risk and Uncertainty*, 2010, pp. 1–25.

Dercon, S., "Growth and shocks: evidence from rural Ethiopia," *Journal of Development Economics*, 2004, *74* (2), 309–329.

——**, J. Gunning, and A. Zeitlin**, "The demand for insurance under limited credibility: evidence from Kenya," 2011.

Eckel, Catherine C and Philip J Grossman, "Sex differences and statistical stereotyping in attitudes toward financial risk," *Evolution and Human Behavior*, 2002, *23* (4), 281–295.

Eckel, C.C. and P.J. Grossman, "Forecasting risk attitudes: An experimental study using actual and forecast gamble choices," *Journal of Economic Behavior & Organization*, 2008, *68* (1), 1–17.

Fafchamps, M., *Rural Poverty, Risk and Development*, Edward Elgar Publishing, 2003.

Fafchamps, Marcel, "Vulnerability, Risk Management, and Agricultural Development," *CSAE Working Paper Series*, 2009.

Falco, P., "Does risk matter for occupational choices? Experimental Evidence from an African labour market," *Manuscript submitted for publication*, 2012.

Falk, A. and J.J. Heckman, "Lab experiments are a major source of knowledge in the social sciences," *Science*, 2009, *326* (5952), 535.

Falk, Armin and Markus Knell, "Choosing the Joneses: Endogenous goals and reference standards," *The Scandinavian Journal of Economics*, 2004, *106* (3), 417–435.

Farber, H.S., "Reference-dependent preferences and labor supply: The case of New York City taxi drivers," *The American Economic Review*, 2008, *98* (3), 1069–1082.

Fechner, G., "Elements of psychophysics. Vol. I.," 1966.

Giné, X., HB Lilleor, RM Townsend, and J. Vickery, "Weather Insurance in Semi-arid India," 2005. Paper presented at the Annual Bank Conference on Development Economics, May 23Ű-24, Amsterdam.

____ , R. Townsend, and J. Vickery, "Statistical Analysis of Rainfall Insurance Payouts in Southern India," *American Journal of Agricultural Economics*, 2007, *89* (5), 1248–1254.

____ , ____ , and ____ , "Patterns of rainfall insurance participation in rural India," *The World Bank Economic Review*, 2008, *22* (3), 539.

Gonzalez, R. and G. Wu, "On the shape of the probability weighting function," *Cognitive Psychology*, 1999, *38* (1), 129–166.

Harbaugh, W.T., K. Krause, and L. Vesterlund, "Risk attitudes of children and adults: Choices over small and large probability gains and losses," *Experimental Economics*, 2002, *5* (1), 53–84.

Harless, D.W. and C.F. Camerer, "The predictive utility of generalized expected utility theories," *Econometrica: Journal of the Econometric Society*, 1994, pp. 1251–1289.

Harrison, G.W. and E.E. Rutström, "Risk aversion in the laboratory," in J. C. Cox and G. W. Harrison, eds., *Research in Experimental Economics*, Vol. 12, Emerald Group Publishing Limited, 2008, pp. 41–196.

____ and ____ , "Expected utility theory and prospect theory: One wedding and a decent funeral," *Experimental Economics*, 2009, *12* (2), 133–158.

____ and J.A. List, "Field experiments," *Journal of Economic Literature*, 2004, *42* (4), 1009–1055.

____ , ____ , and C. Towe, "Naturally occurring preferences and exogenous laboratory experiments: A case study of risk aversion," *Econometrica*, 2007, *75* (2), 433–458.

——— **, M.I. Lau, and E.E. Rutström**, "Theory, experimental design and econometrics are complementary (and so are lab and field experiments)," *The Methods of Modern Experimental Economics*, 2011.

——— **, S.J. Humphrey, and A. Verschoor**, "Choice under Uncertainty: Evidence from Ethiopia, India and Uganda," *The Economic Journal*, 2010, *120* (543), 80–104.

Hazell, P., "The appropriate role of agricultural insurance in developing countries," *Journal of International Development*, 1992, *4* (6), 567–581.

Hey, J.D., "Experimental investigations of errors in decision making under risk," *European Economic Review*, 1995, *39* (3-4), 633–640.

——— **and C. Orme**, "Investigating generalizations of expected utility theory using experimental data," *Econometrica: Journal of the Econometric Society*, 1994, pp. 1291–1326.

——— **and E. Carbone**, "Stochastic choice with deterministic preferences: An experimental investigation," *Economics Letters*, 1995, *47* (2), 161–167.

Hill, R. V. and M. Nobles, "Flexible insurance for heterogenous farmers: results from a small scale pilot in Ethiopia," *IFPRI discussion paper 01092*, 2011.

Hill, RV, A Viceisza, and J Deustua-Rossel, "The welfare and behavioral impact of insurance provision in rural Ethiopia," *Report submitted to USAID*, 2009.

Hill, R.V., J. Hoddinott, and N. Kumar, "Adoption of weather index insurance: Learning from willingness to pay among a panel of households in rural Ethiopia," *IFPRI discussion paper 01088*, 2011.

Holt, C.A. and S.K. Laury, "Risk aversion and incentive effects," *American Economic Review*, 2002, *92* (5), 1644–1655.

Humphrey, S.J. and A. Verschoor, "Decision-making under risk among small farmers in East Uganda," *Journal of African Economies*, 2004, *13* (1), 44.

——— **and** ——— , "The probability weighting function: experimental evidence from Uganda, India and Ethiopia," *Economics Letters*, 2004, *84* (3), 419–425.

Kahneman, D. and A. Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 1979, *47* (2), 263–292.

Koszegi, B. and M. Rabin, "A model of reference-dependent preferences," *The Quarterly Journal of Economics*, 2006, *121* (4), 1133–1165.

——— **and** ——— , "Reference-dependent risk attitudes," *The American Economic Review*, 2007, *97* (4), 1047–1073.

Lee, J., "The effect of the background risk in a simple chance improving decision model," *Journal of Risk and Uncertainty*, 2008, *36* (1), 19–41.

Liu, Elaine M, "Time to change what to sow: Risk preferences and technology adoption decisions of cotton farmers in China," *Review of Economics and Statistics*, 2008.

Loomes, G., "Modelling the stochastic component of behaviour in experiments: Some issues for the interpretation of data," *Experimental Economics*, 2005, *8* (4), 301–323.

—— **, P.G. Moffatt, and R. Sugden**, "A microeconometric test of alternative stochastic theories of risky choice," *Journal of risk and Uncertainty*, 2002, *24* (2), 103–130.

Luce, R.D., "Individual choice behavior.," 1959.

Lusk, J.L. and K.H. Coble, "Risk aversion in the presence of background risk: Evidence from an economic experiment," *Research in Experimental Economics*, 2008, *12*, 315–340.

Lybbert, Travis J, Francisco B Galarza, John McPeak, Christopher B Barrett, Stephen R Boucher, Michael R Carter, Sommarat Chantarat, Aziz Fadlaoui, and Andrew Mude, "Dynamic Field Experiments in Development Economics: Risk Valuation in Morocco, Kenya, and Peru," *Agricultural and Resource Economics Review*, 2010, *39*, 1–17.

McFadden, D., "Conditional Logit Analysis of Qualitative Choice Behavior," *Frontiers in Econometrics*, 1974, pp. 105 –142.

Mosley, P. and A. Verschoor, "Risk attitudes and the 'vicious circle of poverty'," *The European Journal of Development Research*, 2005, *17* (1), 59–88.

Munro, Alistair and Robert Sugden, "On the theory of reference-dependent preferences," *Journal of Economic Behavior & Organization*, 2003, *50* (4), 407–428.

Oehlert, G.W., "A note on the delta method," *American Statistician*, 1992, *46* (1), 27–29.

Padula, Giovanna and Bruno Busacca, "The asymmetric impact of price-attribute performance on overall price evaluation," *International Journal of Service Industry Management*, 2005, *16* (1), 28–54.

Prelec, D., "The probability weighting function," *Econometrica*, 1998, *66* (3), 497–527.

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

—— , *Generalized expected utility theory: The rank-dependent model*, Springer, 1993.

Rabin, M., "Risk aversion and expected-utility theory: A calibration theorem," *Econometrica*, 2000, *68* (5), 1281–1292.

Rosenzweig, M.R. and H.P. Binswanger, "Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments," *The Economic Journal*, 1993, *103* (416), 56–78.

Saha, A., "Expo-power utility: A'flexible' form for absolute and relative risk aversion," *American Journal of Agricultural Economics*, 1993, pp. 905–913.

Schmidt, Ulrich, "Efficient risk-sharing and the dual theory of choice under risk," *The Journal of Risk and Insurance*, 1999, *66* (4), 597–608.

—— **and John D Hey**, "Are preference reversals errors? An experimental investigation," *Journal of Risk and Uncertainty*, 2004, *29* (3), 207–218.

Segal, Uzi and Avia Spivak, "First order versus second order risk aversion," *Journal of Economic Theory*, 1990, *51* (1), 111–125.

Spears, Dean, "Poverty and probability: aspiration and aversion to compound lotteries in El Salvador and India," *Experimental Economics*, 2012, pp. 1–22.

Sprenger, C., "An Endowment Effect for Risk: Experimental Tests of Stochastic Reference Points," *Department of Economics Working Paper, University of California at San Diego*, 2010.

Starmer, C., "Cycling with rules of thumb: An experimental test for a new form of non-transitive behaviour," *Theory and Decision*, 1999, *46* (2), 139–157.

——, "Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk," *Journal of Economic Literature*, 2000, *38* (2), 332–382.

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

Tanaka, Y. and A. Munro, "Regional Variation in Risk and Time Preferences: Evidence from a Large-Scale Field Experiment in Rural Uganda," *National Graduate Institute for Policy Studies Working Paper, Tokyo*, 2012.

Townsend, Robert M., "Risk and Insurance in Village India," *Econometrica*, 1994, *62* (3), 539–591.

Train, K., *Discrete choice methods with simulation*, second ed., Cambridge Univ Pr, 2009.

Tversky, A. and D. Kahneman, "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and uncertainty*, 1992, *5* (4), 297–323.

von Neumann, John and Oskar Morgenstern, *The Theory of Games and Economic Behavior* 1944.

Vuong, Q.H., "Likelihood ratio tests for model selection and non-nested hypotheses," *Econometrica: Journal of the Econometric Society*, 1989, *57* (2), 307–333.

Wakker, P., R. Thaler, and A. Tversky, "Probabilistic insurance," *Journal of Risk and Uncertainty*, 1997, *15* (1), 7–28.

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

Wilcox, N.T., "Stochastic models for binary discrete choice under risk: A critical primer and econometric comparison," 2008.

——, "A Comparison of Three Probabilistic Models of Binary Discrete Choice Under Risk," *Working Paper*, 2010.

APPENDIX A: DERIVATION OF THE STRUCTURAL ESTIMATION MODEL

The structural estimation methodology used to estimate model parameters – described in Section 5.1 – is of the form of the "single-agent stochastic choice" model pioneered by Camerer and Ho (1994). It makes the traditional representative agent assumption that each subject in the experiment has the same preferences – that is, each participant's choices under risk are assumed to be generated by the same decision model with the same preference parameter values.[105,106] However, with this assumption, a stochastic component is required to explain why different participants make different choices in a particular decision problem. In order to incorporate and model this stochastic element, I use the logit discrete choice model, which was developed by Luce (1959) and McFadden (1974). Train (2009) notes that the logit model is the most widely used discrete choice model. In addition, I interpret this model within the framework of the "white noise" error approach, as is common in studies of decision-making under risk (for example, Hey and Orme 1994, Loomes 2005, Harrison et al. 2011).

In this model, the expected utility obtained by a constant relative risk averse decision-maker from alternative i in decision problem d, EU_i^d, represents the individual's true preferences, and is computed using Equation (2) in conjunction with the utility function being considered.[107] Since I assume that all subjects have the same true preferences, each subject obtains the same true expected utility, EU_i^d. Thus, in decision problem d, alternative i is truly preferred to alternative j, where $i, j \in \{1, \ldots, 6\}$ and $i \neq j$, by any individual if and only if $EU_i^d > EU_j^d$. It follows that if choices perfectly reflect true preferences, any participant would choose alternative i in decision problem d if and only if $EU_i^d > EU_j^d \; \forall j \neq i \in \{1, \ldots, 6\}$.

However, I assume that some noise is liable to enter into the decision-making process, so that the actual choice made by the individual in the decision problem is not based solely on the true expected utility (EU_i^d), but on *perceived* expected utility, where the perceived expected utility obtained by individual a from

105 While the single preference assumption is restrictive, it is required for the sake of parsimony and estimation, since each participant in the Ethiopian experiment only made a small number of choices (Camerer and Ho 1994). Further, Camerer and Ho (1994) note that the "single-agent stochastic choice" approach provides a simple and tractable method for pooling data from different individuals and decision problems, when estimates at the individual level may be impossible to obtain or have very low power (that is, when a within-subject analysis is not feasible). Additionally, by separately analyzing the choices in the benchmark problem and insurance problems, I allow the decision models (and corresponding preference parameters) generating the choices in these problems to vary.

106 Hey (1995) notes that the representative agent in these models can be regarded as the average subject involved in the experiment, and thus parameter estimates of the decision models can be interpreted as the parameter values for the average subject. Thus, the single-agent model considered here – which has also been used by Andersen et al. (2006a) and Harrison et al. (2011) in the estimation of risk preferences – is in line with the numerous economic theories that assume a single representative agent, and therefore the single parameter estimates obtained from this model are important as empirical inputs to such theorizing. Further, it is useful to know which functional form and corresponding single parameter values best fit the average subject (Camerer and Ho 1994).

107 In this description, I assume EUT-CRRA preferences for illustrative purposes; the structural model is easily extended to other decision theories.

alternative i is defined as $EU_i^d + \varepsilon_{ai}^d$. ε_{ai}^d is the (logit) error term associated with alternative i for individual a, and causes reported preferences (that is, observed choices) to deviate from true preferences. Thus, while participants have the same true expected utility and true preferences, their perceived expected utility and perceived preferences differ, since the logit error term ε_{ai}^d differs between individuals – this allows for heterogeneity in the observed choices of subjects. This "white noise" error may arise due to carelessness on the part of individuals when evaluating alternatives in a decision problem, due to imperfect access to their true utility function, or as a result of subjects misunderstanding the nature of the decision problem (Hey 1995, Loomes 2005).

Taking the noise parameter into account, subject a chooses alternative i in decision problem d if and only if $EU_i^d + \varepsilon_{ai}^d > EU_j^d + \varepsilon_{aj}^d \; \forall j \neq i \in \{1, \ldots, 6\}$. Further, the logit model assumes that each ε_{ai}^d is independently and identically distributed with Type I Extreme Value (or Gumbel) distribution[108]; this implies that the errors are independent across alternatives (in the same decision problem and in different decision problems) and across individuals (Train 2009).[109,110]

Following Train (2009), I note that the probability that decision-maker a chooses alternative i out of the six alternatives in decision problem d is given by:

$$\begin{aligned} P_{ai}^d &= \text{Prob}(EU_i^d + \varepsilon_{ai}^d > EU_j^d + \varepsilon_{aj}^d \forall j \neq i \in \{1, \ldots, 6\}) && (25) \\ &= \text{Prob}(\varepsilon_{aj}^d < \varepsilon_{ai}^d + EU_i^d - EU_j^d \forall j \neq i \in \{1, \ldots, 6\}) && (26) \end{aligned}$$

Given the assumption that the errors are independently and identically distributed with Type I Extreme Value distribution, Train (2009) shows that this expression can be manipulated to obtain the logit choice probability:

$$P_{ai}^d = \frac{\exp(EU_i^d)}{\sum\limits_{j=1}^{6} \exp(EU_j^d)} \tag{27}$$

It is important to note that since the true preferences, represented by the expected utility in this case, are the same for all individuals in the representative agent model, the probability of choosing an alternative i in decision problem d is the same for all subjects. Thus, $P_{ai}^d = P_i^d$, and the probability only varies between alternatives and decision problems, and is the same for all individuals for a particular alternative in a given decision problem.

108 The variance of this distribution is $\pi^2/6$.

109 Cameron and Trivedi (2005) note that while this independence assumption makes the model analytically tractable, it is fairly restrictive, and implies the assumption of independence of irrelevant alternatives (IIA). However, the assumption is not as restrictive as it might at first seem – indeed, it can be considered a natural outcome of a well specified model (Train 2009). Train (2009) points out that even if the errors are correlated over alternatives, the logit model – with representative utility – provides an appropriate approximation of decision-making under risk, since violations of the logit assumptions do not seem to have a significant effect when estimating average preferences.

110 It is important to note that the logit discrete choice model described here is very similar to the multinomial logit model. Both models assume that the error term has Type I Extreme Value distribution and provide similar expressions for the probabilities of observing different choices (outcomes).

If I follow the notation in Section 5.1 and consider $\exp(EU_i^d) = eu_i^d$, then I obtain the expression for the latent index ∇EU_i^d (described in Equation (11)), which is interpreted as the probability of an individual choosing option i in decision problem d:

$$P_i^d = \nabla EU_i^d = \frac{eu_i^d}{\sum\limits_{j=1}^{6} eu_j^d} \tag{28}$$

I can then use this latent index, which represents the probability of observing a particular choice, to construct the log-likelihood function for the sample of choices in the decision problems, as described in Section 5.1; the log-likelihood function can then be maximized to obtain maximum likelihood estimates of the model parameters (in this case, r).

Thus, this represents a model of "stochastic choice with deterministic preferences", since the true preferences are deterministic (given by EUT-CRRA in this case) but the observed choices are stochastic due to the presence of "white noise" error (Hey and Carbone 1995); that is, observed choices can differ from true preferences. This model incorporates noise in decision-making (through the inclusion of the error term ε_{ai}^d), though the variance of the error term ε_{ai}^d (and thus the magnitude of noise in decision-making) is fixed.[111] This logit model and maximum likelihood estimation technique has been used to analyze data from an experiment with Ordered Lottery Selection design conducted by Harrison and Rutström (2008) and involving university students in the United States; it has not, however, to my knowledge, been used extensively to analyze Ordered Lottery Selection data from experiments conducted in developing countries.

[111] Thus, it cannot be used to estimate the extent of randomness in experimental decisions.

APPENDIX B: DERIVATION OF THE STOCHASTIC ERROR MODELS

Luce error specification: In deriving the logit model used for the structural estimations, described in Appendix A, I assumed that the error term ε_{ai}^{d} is distributed Type I Extreme Value with variance $\pi^2/6$. Setting (or normalizing) the variance to be $\pi^2/6$ (which is a direct consequence of the distributional assumptions made for the error term) is equivalent to normalizing the scale of utility (Train 2009).

More generally, if EUT-CRRA is assumed, the perceived expected utility referred to in Appendix A can be considered to be $EU_i^d + \varepsilon_{ai}^{*d}$, where the variance of the error term is now $\mu^2 \times (\pi^2/6)$; that is, I no longer consider the variance to be fixed at $\pi^2/6$, but allow it to be any number, re-expressed as a multiple of $\pi^2/6$. Perceived expected utility can be divided by μ without changing behaviour, since the scale of utility is irrelevant to behaviour (Train 2009). Thus, perceived expected utility now becomes $EU_i^d/\mu + \varepsilon_{ai}^d$ where $\varepsilon_{ai}^d = \varepsilon_{ai}^{*d}/\mu$. Train (2009) notes that in this transformed model the error term ε_{ai}^d once again has variance $\pi^2/6$ and, as shown in Appendix A, the probability of individual a choosing alternative i in decision problem d (which is the same for all individuals) is:

$$P_{ai}^d = P_i^d = \frac{\exp(EU_i^d/\mu)}{\sum\limits_{j=1}^{6} \exp(EU_j^d/\mu)} \tag{29}$$

If I consider $\exp(EU_i^d/\mu) = eu_{iLuce}^d$, then I obtain the expression for the latent index ∇EU_{iLuce}^d for the Luce error specification (described in Equation (15)), which is interpreted as the probability of an individual choosing i in decision problem d (that is, $\nabla EU_{iLuce}^d = P_i^d$). This latent index can be used to construct the log-likelihood function – for the Luce error model – for the sample of choices in the decision problems, as described in Section 5.2; the log-likelihood function can then be maximized to obtain maximum likelihood estimates of the preference parameters (in this case, r) and the noise parameter (μ).

Thus, the parameter μ reflects the variance of the error term, as its value determines the size (or scale) of the error variance. Train (2009) refers to μ as the scale parameter, as it scales the expected utility to reflect the variance of the error. From Equation (29), it can be inferred that as $\mu \to \infty$, the index $\nabla EU_{iLuce}^d \to 1/6$ – thus, the probability of choosing each of the six options tends to $1/6$, and the choice essentially becomes random, or completely removed from true preferences (that is, true expected utility). Meanwhile, as $\mu \to 0$, the specification collapses to the deterministic EUT model, where the choice is strictly determined by the expected utilities of the different lotteries and there are no stochastic errors in the decision-making process (Harrison and Rutström 2008). When $\mu = 1$, this specification collapses to the structural estimation specification described in Section 5.1, where the variance of the error is fixed at $\pi^2/6$.

Since the parameter μ determines the extent to which choice probabilities depend on true preferences, I follow Harrison and Rutström (2008) and Harrison et al. (2011) and interpret μ as the structural "noise" parameter in the Luce error model – this parameter reflects the magnitude of the error in the decision-making process. In other words, it reflects the magnitude of the deviation of observed choices from true preferences (given by the expected utilities of the lotteries), with larger values reflecting greater deviation and greater importance of the error, as opposed to true preferences, in determining choice probabilities (and thus observed choices).

Fechner error specification: With only two alternatives in each decision, in the Fechner error framework, the true preferences of an individual a over the two alternatives (R and L) in a lottery pair p in decision problem d can be represented by the difference in expected utilities of the two alternatives (Loomes 2005). That is, true preferences can be indicated by the net expected utility, $EU^d_{apR} - EU^d_{apL}$, and if the subject does not make any mistakes, the right lottery is chosen if and only if $EU^d_{apR} - EU^d_{apL} > 0$; otherwise the left lottery is chosen.

However, if the subject makes "white noise" errors of the form described in Appendix A, then the observed choices could deviate from true preferences. In this case, choices are made on the basis of *perceived* net expected utility, $EU^d_{apR} - EU^d_{apL} + \varepsilon^{*d}_{ap}$, where ε^{*d}_{ap} represents the "white noise" error in this model. The right lottery is now chosen if and only if perceived net expected utility is positive. Taking these errors into account, the probability that subject a chooses the right lottery in lottery pair p of decision problem d is given by:

$$P^d_{aR} = \text{Prob}(EU^d_{apR} - EU^d_{apL} + \varepsilon^{*d}_{ap} > 0) \quad (30)$$

$$= \text{Prob}(\varepsilon^{*d}_{ap} > -(EU^d_{apR} - EU^d_{apL})) \quad (31)$$

Following Harrison et al. (2007), I assume that ε^{*d}_{ap} is independently and identically normally distributed with mean 0 and standard deviation μ; this implies that the errors are independent across lottery pairs (in the same decision problem and in different decision problems) and across individuals. Given this distribution for ε^{*d}_{ap}, $\varepsilon^{*d}_{ap}/\mu$ is normally distributed with mean 0 and standard deviation 1 (standard normal distribution). Thus, it follows from Equation (31) that:

$$P^d_{aR} = \text{Prob}(\varepsilon^{*d}_{ap}/\mu > -(EU^d_{apR} - EU^d_{apL})/\mu) \quad (32)$$

$$= 1 - \Phi\left(\frac{-(EU^d_{apR} - EU^d_{apL})}{\mu}\right) \quad (33)$$

$$= \Phi\left(\frac{EU^d_{apR} - EU^d_{apL}}{\mu}\right) \quad (34)$$

Thus, the probability of choosing the left lottery is given by:

$$P^d_{aL} = 1 - \Phi\left(\frac{EU^d_{apR} - EU^d_{apL}}{\mu}\right) = \Phi\left(\frac{-(EU^d_{apR} - EU^d_{apL})}{\mu}\right) \quad (35)$$

where Φ is the standard normal cumulative distribution function, and I utilize the symmetry (around 0) of the standard normal distribution in the above manipulations.[112] It is important to note that since the true preferences, represented by expected utility in this case, are the same for all individuals in the representative agent model (as explained in Appendix A), the above expressions for the probabilities of choosing the right and left lotteries in pair p of decision problem d are the same for all subjects. Thus, $P^d_{aR} = P^d_R$ and $P^d_{aL} = P^d_L$, and the probability only varies between lottery pairs and decision problems.

Further, if I consider the latent index $\nabla EU^d_{pFech} = (EU^d_{apR} - EU^d_{apL})/\mu$, then the expressions for the probabilities of the right and left lotteries being chosen can be obtained in terms of the latent index, as described in Equations (17) and (18) (in Section 5.2). These probabilities can be used to construct the log-likelihood function – for the Fechner error model – for the sample of choices in the decision problems, as described in Section 5.2; the log-likelihood function can then be maximized to obtain maximum likelihood estimates of the preference parameters (in this case, r) and the noise parameter (μ).

As $\mu \to 0$, the error specification collapses to the deterministic EUT model, where the choice in a lottery pair is strictly determined by the true expected utilities of the two lotteries and the option with greater expected utility is always chosen. As $\mu \to \infty$, the choice essentially becomes random and does not depend on the true preferences (that is, on the difference between the expected utilities of the two lotteries). Therefore, similar to the interpretation used in the Luce error model, μ (which is the standard deviation of the error term) is interpreted as the structural noise (or error) parameter in the Fechner model, with larger values of μ implying greater stochastic errors in the decision-making process (that is, greater deviation of observed choices from true preferences).[113]

[112] As noted by Harrison and Rutström (2008), the logistic cumulative distribution function can also be used, instead of the standard normal cumulative distribution function, to link observed choices to true preferences. In this case, the error ε^{*d}_{ap} is assumed to have a logistic distribution with mean 0 and variance $\mu^2 \times (\pi^2/3)$, and the transformed error $\varepsilon^{*d}_{ap}/\mu$ then has a logistic distribution with mean 0 and variance $\pi^2/3$.

[113] It is important to note that some experimental studies (for example, Harrison and Rutström 2008), including Chapter 3 of this thesis, use versions of this Fechner structural model with the error variance fixed, analogous to the model outlined in Appendix A. In this case, using the standard normal cumulative distribution link function, the error ε^{*d}_{ap} is assumed to be normally distributed with mean 0 and standard deviation 1; similarly, using the logistic cumulative distribution link function, the error ε^{*d}_{ap} is assumed to have a logistic distribution with mean 0 and variance $\pi^2/3$. In either case, the variance of the error term – and thus the magnitude of noise in the decision-making process – is fixed and cannot be estimated, though the model still incorporates (a fixed degree of) randomness in choice and stochastic errors in the decision-making process.

ASSET INTEGRATION AND THE DETERMINANTS OF RISK PREFERENCES: EVIDENCE FROM RURAL ETHIOPIA

GAUTAM KALANI*

Abstract

In this chapter, I analyze combined experimental and survey data from rural Ethiopia in order to evaluate the determinants of risk preferences of inhabitants as well as assess the degree of asset integration in experimental decisions. I focus on the impact of two important economic factors – wealth and background risk – on risk attitudes, and account for the potential endogeneity associated with these two variables by using an instrumental variable strategy in conjunction with interval regression, Limited Information Maximum Likelihood and structural maximum likelihood estimation techniques. Analyzing both expected utility theory and non-expected utility theory decision models, I find some evidence that household wealth negatively affects both risk aversion and loss aversion. However, the results also indicate that independent background risk has no significant impact on the risk preferences of individuals. Further, I find evidence of narrow framing, as opposed to asset integration, suggesting that participants isolate decisions in the experimental lotteries and make choices by considering only the prospective gains and losses associated with the current decision, independent of outside wealth.

JEL codes: C36, C91, D01, D03, D81, O16.

*DPhil student: Department of Economics, Centre for the Study of African Economies (CSAE) and Hertford College, University of Oxford (**gautam.kalani@economics.ox.ac.uk**). This paper is part of the work towards my DPhil thesis, which is supervised by Professor Marcel Fafchamps; the work would not have been possible without his very generous assistance. Additionally, I am grateful to Daniel Clarke, Simon Quinn and Riccardo Lucchetti, who offered useful suggestions throughout the project.

1 INTRODUCTION

In order to obtain a comprehensive understanding of decision-making under risk, it is crucial to evaluate the determinants of risk attitudes, in addition to identifying and estimating the risk preferences themselves. Further, recent experimental studies in developing countries (for example, Yesuf and Bluffstone 2009, Tanaka and Munro 2012) have found that risk attitudes are inextricably tied to – and determined by – demographic factors and household socioeconomic circumstances. In this chapter, I analyze combined experimental and survey data from rural Ethiopia in order to evaluate the determinants of risk preferences of inhabitants as well as assess the degree of asset integration in experimental decisions.

In particular, I focus on jointly establishing the impact of two economic factors – wealth and background risk – on risk attitudes.[1] As noted by Guiso and Paiella (2008), the relationship between risk preferences and wealth is central to numerous fields of economics, and is especially important for predicting economic reactions under risk and uncertainty. Rosenzweig and Binswanger (1993) and Wik et al. (2004) note that wealth is a crucial determinant of risk preferences, and find that wealthier households exhibit significantly lower aversion to risk, possibly due to the cognitive effects of long-run material success that reduce risk aversion.[2] However, as noted by Liu (2008), the empirical evidence on the relationship between wealth and risk preferences is far from conclusive. While the above-mentioned studies find a negative relationship between wealth and risk aversion, some studies find no significant relationship between the two (for example, Binswanger 1980, Mosley and Verschoor 2005, Falco 2012), and still others find a positive association between wealth and risk aversion (for example, Nielsen 2001). Indeed, Banerjee (2000) suggests

[1] As done in Chapter 1, I follow the nomenclature used by Harrison and Rutström (2008b) and refer to the properties (or parameters) of the utility function as risk attitudes or risk preferences.

[2] Further, Yesuf and Bluffstone (2009) find that wealth accumulation affects both risk exposure and risk preferences, and represents a critical long-run link between risk aversion, path dependence and poverty.

that the poor may be more or less risk averse than the non-poor, depending on whether the effect of "vulnerability" (losses cause them too much pain) or that of "desperation" (they have little to lose) dominates; therefore, empirically distinguishing between these two theoretical possibilities is important, as the two alternatives have very different policy implications.

Most studies that analyze risk-taking focus on individuals' behaviour when faced with a single risky decision – for example, the decision to buy insurance or the choice in an experimental decision problem (Lusk and Coble 2008). However, in reality, individuals are simultaneously faced with a variety of risks rather than a single risk, and decisions in relation to a particular risk can be influenced by other exogenous, uncorrelated "background" risks, which usually cannot be avoided or insured against (Guiso and Paiella 2008, Lusk and Coble 2008). Therefore, it is important to account for these independent background risks in order to accurately identify the determinants of risk preferences and obtain a more complete understanding of decision-making under risk.

Gollier and Pratt (1996) derive conditions, which they label "risk vulnerability", under which the addition of an independent background risk would cause an increase in aversion to other risks. They note that familiar decreasing absolute risk aversion (DARA) utility functions – including constant relative risk aversion (CRRA) – are risk vulnerable, and postulate that an increase in background risk would lead to greater risk aversion in relation to other risky decisions or situations if these preferences are assumed. On the other hand, Diamond (1984) and Quiggin (2003) investigate conditions under which aversion to a particular risk would be reduced by the addition of independent background risk. Indeed, Lusk and Coble (2008) note that determining the sign and magnitude of impact, if any, of background risk on risk aversion is ultimately an empirical issue, as theoretical expositions do not provide a clear prediction of this relationship. However, there are relatively few empirical studies on the subject, and results from these studies are inconclusive (Lee 2008, Lusk and Coble 2008). Guiso and Paiella (2008), Harrison et al. (2007) and Lee (2008) find that background risk significantly increases risk aversion. On the other hand, Masson and Arrondel (1996) find that the ownership of risky assets rises with earnings risk, while Alessie et al. (2002) and Lusk and Coble (2008) find that background risk does not have any significant impact on risk aversion. Indeed, Lusk and Coble (2008) and Tanaka and Munro (2012) stress the importance of further experimental research on the impact of background risk on risk preferences; this importance is magnified for low-income rural populations, who are exposed to substantial uninsured risk from a multitude of sources.

Therefore, questions regarding the direction and magnitude of the effect of both wealth and background risk on risk aversion remain unresolved, despite the potential importance of these factors in explaining the risk-taking and investment decisions of individuals, and thus for understanding the economic circumstances of inhabitants of developing countries. However, economic factors, such as wealth and background risk, are likely to be affected by risk preferences and thus endogenously determined due to simultaneity (reverse causality), thereby making it difficult to establish the direction of causality (Guiso and Paiella 2008). For example, significant aversion to risk often causes the rural poor to forego investment activities that have high risk but also higher expected returns (such as the adoption of high-yield crop varieties) – they instead undertake low-risk, low-return investment and production strategies (such as subsistence agriculture) which hinder technology adoption and wealth accumulation, leading to persistent, long-term poverty and poverty traps (Banerjee and Newman 1993, Rosenzweig and Binswanger 1993, Mosley and Verschoor 2005).[3] On the other hand, as noted by Guiso and Paiella (2008), it is also possible that risk averse individuals are more prudent, thus compressing current consumption in order to save and accumulate more asset wealth. Liu (2008) also finds that both loss aversion and risk aversion affect technology adoption decisions, while Hill et al. (2011) find that risk attitudes are an important determinant of insurance purchase decisions

[3] Additionally, Mosley and Verschoor (2005) note that this is not necessarily evidence of irrationality, but may instead be a result of a finely balanced survival algorithm.

– these crucial economic decisions have a significant bearing on both household wealth and exposure to background risk.

In order to deal with this issue of endogeneity and obtain consistent estimates of the causal impact of wealth and background risk on risk preferences, I employ an instrumental variable (IV) strategy, using mainly village-level variables – distance to the nearest town, village size and the inter-temporal standard deviation of annual village-level rainfall – interacted with the gender of the respondent as instrumental variables for wealth and background risk. Gender is a demographic characteristic that is not endogenously determined by risk preferences, and while many studies find that risk preferences vary by gender (for example, Eckel and Grossman 2008, Guiso and Paiella 2008) – as a result of which gender is included as a determinant of risk attitudes – there is little reason to suspect that this correlation varies by village and village-level factors such as distance to the nearest town and village size. Therefore, these interaction terms are not expected to directly affect risk attitudes and thus can be excluded as determinants of risk preferences (that is, they are expected to satisfy the exclusion restriction). Further, village characteristics such as distance to the nearest town and village size are expected to affect household wealth, while the inter-temporal standard deviation of village-level rainfall is expected to impact background risk. Additionally, the influence of these village-level characteristics on household wealth and background risk is expected to vary by the gender of the respondent, who in most cases is the household head, and thus in control of household finances. This is because, as noted by Powell and Ansic (1997), male and female respondents likely account for these village-level factors differently in household financial decision-making, implying a differential impact (by gender) of these factors on household wealth and background risk. Thus, these interaction terms are expected to impact household wealth and background risk, and influence risk preferences only through their effect on these variables.

To estimate the IV models, I use both two-step and Limited Information Maximum Likelihood (LIML) interval estimators, as described by Bettin and

Lucchetti (2010), as well as a two-step version of the structural maximum likelihood estimator detailed by Harrison and Rutström (2008b); these estimations enable me to establish the causal impact of wealth and background risk on risk preferences, which are not directly observed but are inferred from experimental choices.[4] To my knowledge, these estimators have not previously been used in the analysis of the determinants of risk preferences.

Analyzing choice data from Binswanger-style lotteries in the Ethiopian Rural Household Survey (ERHS) and in a laboratory experiment conducted in rural Ethiopia (which was also analyzed in Chapter 1) within an expected utility theory (EUT) framework, I find that background risk does not significantly affect risk aversion; on the other hand, there is some evidence that wealthier respondents exhibit less risk aversion in these decision problems. Further, analyzing data from the insurance decision problems of the Ethiopian experiment within a non-EUT framework, I find that wealth has a negative impact on both risk aversion and loss aversion, while once again background risk has no significant role in determining risk preferences.

An important aim of this study, in addition to establishing the determinants of risk preferences, is to test the asset integration hypothesis; that is, to evaluate whether subjects integrate their household wealth with potential lottery earnings when making choices in the experimental decision problems. Heinemann (2008) notes that although many economic theories of decision-making under risk assume agents fully integrate income from all sources in every decision, there is a large body of theoretical and empirical evidence that casts serious doubt on such behaviour. In particular, Arrow (1971) notes that while maximizing expected utility from total wealth implies almost risk-neutral behaviour with respect to decisions involving small monetary amounts (relative to wealth), most experimental studies involving such amounts find that participants

[4] This chapter therefore also contributes to the debate on the endogeneity of risk attitudes and the direction of causality between economic factors and risk preferences – Falco (2012) notes that this is an important avenue for research with crucial implications for both economic policy and theoretical modeling.

exhibit substantial risk aversion. Conversely, as Rabin (2000) indicates, even moderate risk aversion over the small stakes in laboratory experiments implies inconceivably high risk aversion over large stakes, if individuals are assumed to maximize expected utility from total wealth.

In fact, there is no consensus in the literature regarding the arguments of the utility function, and Harrison et al. (2007) believe that researchers are free to define utility over virtually anything. Within the framework of expected utility theory (EUT), utility is generally considered to be defined either over total wealth (expected utility from wealth), which involves the full integration of assets and income from all sources, or only over the prospective gains and losses associated with the current decision (expected utility from income), independent of outside wealth (Heinemann 2008). The former phenomenon has been referred to as full "asset integration", while the latter has been labelled as "narrow framing". Harrison et al. (2007) note that even though theoretical predictions are ambiguous and most studies impose one or the other assumption *a priori*, it is possible – and important – to distinguish empirically between expected utility of wealth and income.

Thus, following the maximum likelihood estimation strategy used by Harrison et al. (2007) and Heinemann (2008), I estimate the argument of the utility function in order to determine whether asset integration or narrow framing better describes the decisions of individuals in the experimental problems. In particular, I define the argument of the utility function as $\delta W + x$ (where x is the lottery prize and W is outside wealth), and estimate the parameter δ, which indicates the degree of asset integration (or the extent to which outside wealth is integrated with lottery prizes in experimental decisions). Thus, this flexible specification incorporates both the possibility of narrow framing ($\delta = 0$) and full asset integration ($\delta = 1$), as well as intermediate levels of asset integration – it lets the data determine, without any strong *a priori* assumptions, the degree of asset integration.

Harrison et al. (2007) and Heinemann (2008), however, do not possess information on subjects' outside wealth W, and hence cannot separately identify δ; thus, they can only estimate δW as an additive constant term in the argument of the utility function. In other words, they can test the hypothesis of narrow framing ($\delta W = 0$), but cannot estimate the extent (or degree) of asset integration. Using the information in the ERHS dataset on household wealth, I build on their work by estimating the δ parameter separately, thus obtaining robust and accurate estimates of the degree of asset integration, in addition to testing the narrow framing hypothesis. Further, most studies analyzing the asset integration hypothesis use data from developed country experiments, and the phenomenon of asset integration in the context of experiments conducted in developing countries – whose inhabitants are exposed to substantial risk but rely primarily on informal risk-coping options – has been largely ignored in the empirical literature.

Analyzing the choices of subjects in the Binswanger-style lotteries of the Ethiopian experiment and ERHS within an EUT framework, I find that narrow framing better describes the decisions of respondents, and outside wealth does not enter as an argument of the utility function. This method of accounting for wealth in decision-making under risk – that is, as a potential argument of the utility function – is different from the analyses in most other studies, which consider risk attitudes to be a function of wealth but assume the argument of the utility function to be x. However, correctly identifying the argument of the utility function is a crucial factor for the accurate estimation of risk preferences.

Most studies that analyze the determinants of risk preferences in developing countries do not appropriately account for the endogeneity associated with household and individual economic circumstances (for example, Yesuf and Bluffstone 2009, Tanaka and Munro 2012).[5] This study therefore builds on the work of Tanaka et al. (2010), who combine survey and experimental data from rural Vietnam and use an IV strategy to estimate the impact of wealth on risk preferences. However, the timing of their experiment did not closely match

[5] Further, many studies focus on evaluating the correlates of risk attitudes, and do not aim to establish or estimate the causal impact of economic factors on risk preferences (for example, Liu 2008, Tanaka and Munro 2012).

that of the survey data collection; the experiment was conducted three years after the survey – this represents a considerable gap, and household economic circumstances were likely to have changed significantly in the time between the survey and the experiment. On the other hand, the experimental and survey data analyzed in this chapter were collected only a few months apart. Further, Tanaka et al. (2010) only consider the relationship between wealth and risk attitudes, and do not account for background risk in their analysis; indeed, there is a relative scarcity of experimental studies that jointly estimate the impact of both wealth and background risk on risk attitudes in low-income settings. Lusk and Coble (2008) note that omitting background risk could generate biased estimates of the impact of other factors on risk preferences and thus incorrect inferences when analyzing risk-taking behaviour. Therefore, following the precedent of Guiso and Paiella (2008), I jointly test the impact of both wealth and background risk on risk preferences; further, while they consider background risk at the provincial level and only instrument for wealth (not background risk), I use household-level measures of wealth and background risk and instrument for both in the IV analysis. To my knowledge, this is the first analysis of the determinants of risk preferences in a low-income setting which involves instrumenting for both wealth and background risk.

Further, while most studies analyzing the determinants of risk preferences use a single-parameter EUT framework, I consider both an EUT and non-EUT framework. As found in Chapter 1, non-EUT decision models with multiple preference parameters better describe choices in the insurance problems of the Ethiopian experiment; this is in line with the results of numerous experimental studies. Thus, research on the determinants of risk attitudes using non-EUT decision models is crucial (an observation also made by Tanaka et al. 2010). In particular, an important aim of this chapter is to evaluate the determinants of loss aversion (in addition to risk aversion), given that recent studies have found that loss aversion may be a more important characterization of the behaviour of poor villagers in developing countries than risk aversion, as it plays a larger role in shaping individual preferences and driving decision-making (for example, Fafchamps 2009, Tanaka and Munro 2012). The non-EUT analysis in this chapter therefore builds on the work of Tanaka et al. (2010), who use a prospect theory framework to analyze the relationship between wealth and risk preferences.

Additionally, the combined experimental and survey dataset used in this chapter provides fertile ground to investigate the influence of an extensive set of demographic and socioeconomic factors on risk preferences. Tanaka et al. (2010) indicates that a broad approach such as this is complementary to more targeted, policy-specific studies (such as that of Ashraf et al. 2006), and both types of analyses are essential. This chapter therefore builds on recent work in experimental development economics which uses such combined datasets to study the relationship between experimentally-measured risk attitudes and economic decisions and circumstances in developing countries (for example, Mosley and Verschoor 2005, Yesuf and Bluffstone 2009, Attanasio et al. 2012). However, very few of these experimental studies are linked to detailed survey data that has been collected in multiple rounds over a long period of time; on the other hand, the ERHS is an extensive survey conducted over seven rounds spanning 15 years, thus enabling me to utilize important panel aspects of this long-term survey as well. Therefore, this chapter, in combination with Chapter 1, provides a comprehensive analysis of the behaviour under risk of individuals in rural Ethiopia.

The rest of this chapter is organized as follows. The next section outlines the conceptual framework and testing strategy, while Section 3 describes the dataset and key variables used in the empirical specifications, including the instrumental variables. The results of the empirical analysis are presented and discussed in Section 4. Section 5 provides robustness checks and Section 6 concludes.

2 CONCEPTUAL FRAMEWORK AND TESTING STRATEGY

2.1 Asset integration

In Chapter 1, I considered a constant relative risk aversion (CRRA) utility function of the form:

$$U(x) = \begin{cases} \dfrac{x^{1-r}}{1-r} & \text{if } r \neq 1 \\ \ln(x) & \text{if } r = 1 \end{cases} \tag{1}$$

where x is the lottery prize and r is the coefficient of relative risk aversion: $r < 0$ corresponds to risk loving behaviour, $r = 0$ to risk neutral and $r > 0$ to risk averse.[6,7] Under EUT, the decision-maker weights each possible outcome $k_i \in \{1, \ldots, K\}$ in lottery choice i using the associated probability p_{k_i}, and so expected utility from lottery i (EU_i) is the sum of the probability weighted utility of each outcome in the lottery:

$$EU_i = \sum_{k_i=1}^{K} p_{k_i} U_{k_i} \tag{2}$$

Therefore, by assuming the argument of the utility function to be the lottery prize x, I assume that participants make decisions by maximizing the expected utility of income from the current decision (without integrating outside wealth), as is commonly done in the experimental literature (for example, Tanaka et

[6] The Arrow-Pratt measure of relative risk aversion, or coefficient of relative risk aversion, is defined as $\frac{-xU''(x)}{U'(x)}$.

[7] As done in Chapter 1, the show-up fee (of 5 birr) in the Ethiopian experiment is excluded from the lottery prize x in all estimations involving data from the experiment; this is common practice in empirical analyses of experimental data, and assumes that participants do not integrate their show-up fee with earnings from the experimental decision problems (Harrison and Rutström 2008b). This also implies that in reference-dependent models (such as those involving loss aversion), the reference point does not include the show-up fee – for example, if the initial endowment or the payoff in the risk-free choice option in the decision problem is considered to be the reference point, these values do not include the show-up fee. However, the results remain substantively the same when the show-up fee is included in the measure of the lottery prize x.

al. 2010, Harrison et al. 2010). However, in this chapter, I aim to explicitly test whether narrow framing provides an accurate description of the experimental choices of participants, or whether subjects integrate their wealth from outside the experiment when making decisions in the experimental problems. Therefore, following Cox and Sadiraj (2004) and Heinemann (2008), I no longer assume that x is the only argument of the utility function and define the CRRA utility function as:

$$U(W,x) = \begin{cases} \dfrac{(\delta W + x)^{1-r}}{1-r} & \text{if } r \neq 1 \\ \ln(\delta W + x) & \text{if } r = 1 \end{cases} \tag{3}$$

where W is the outside or initial wealth of the subject. Using this utility function, expected utility from a potential lottery choice i is computed using Equation (2).[8]

Therefore, the argument of the utility function is now $\delta W + x$, rather than only the lottery prize x. δ is a parameter that defines the degree to which subjects integrate their wealth with prospective prizes in the decision problems, with greater values of δ indicating greater integration of outside wealth in the experimental decisions. In other words, δ is the degree of asset integration, and indicates the weight placed on outside wealth in the argument of the utility function (Andersen et al. 2006a). $\delta = 1$ indicates full asset integration, implying that participants fully integrate wealth from all sources (and not just from the current decision) when making choices in risky situations. $\delta = 0$, on the other hand, indicates narrow framing, implying that participants only consider prospective gains from the current decision when making choices.[9] In accordance with the arguments of Cox and Sadiraj (2006) and the methodology of Andersen et al. (2006a), δ is allowed to take on any value between 0 and 1 (both inclusive).[10] The parameter

[8] Note that the show-up fee (of 5 birr) in the experiment is excluded from the measure of outside wealth.

[9] It is important to note that in experiments where participants are allowed to accumulate wealth over consecutive decision problems, there could also be integration of laboratory income from prior decision problems within the experiment – this "local asset integration" is not accounted for in the above-described model. However, this is unlikely to be an issue in the context of the Ethiopian experiment, since participants only played one of the three decision problems they responded to for real money, and this decision problem was randomly selected at the end of the session.

[10] Thus, this specification allows the special cases of perfectly narrow framing and full asset integration.

r is interpreted in a similar manner as before, and defines the curvature of the utility function with respect to $\delta W + x$.

To estimate the parameters of this model, I use a structural model combined with a maximum likelihood estimation technique (similar to that used in Chapter 1), following the estimation strategy pioneered by Camerer and Ho (1994) and detailed in the work of Harrison and Rutström (2008b). This structural estimation methodology – which incorporates a logit discrete choice model – is as follows (a detailed derivation of this model, including the specific assumptions underlying it, is provided in Appendix A of Chapter 1). First, the expected utility EU_i^d from each potential lottery choice $i \in \{1, \ldots, n\}$ in decision problem d (which involves subjects choosing their most preferred of n options)[11] is calculated according to Equation (2), assuming that utility is defined by the CRRA function in Equation (3). The latent index ∇EU_i^d is then calculated as follows:

$$eu_i^d = \exp(EU_i^d) \tag{4}$$

$$\nabla EU_i^d = \frac{eu_i^d}{\sum_{b=1}^{n} eu_b^d} \tag{5}$$

The latent index ∇EU_i^d, based on latent preferences, is in the form of a probability between 0 and 1, and thus can be directly linked to observed choices. Following Harrison and Rutström (2008b), ∇EU_i^d is interpreted as the probability of a subject choosing lottery choice i in decision problem d.[12]

[11] The Binswanger-style lotteries in the Ethiopian experiment (B) and the ERHS involve six and five options, respectively.

[12] It is important to note that this specification incorporates a stochastic element to allow for errors (or noise) in experimental decision-making that cause observed choices to deviate from true preferences. Thus, it represents a model of "stochastic choice with deterministic preferences". However, in this specification, the extent of errors in decision-making is fixed. In reference to the logit discrete choice model described in Appendix A and B of Chapter 1, this implies that the variance of the logit error term is fixed and the scale of utility is normalized; additionally, in reference to the Luce error model described in Chapter 1, this specification assumes that the noise parameter μ, which determines the variance of the logit error term, is equal to 1. Numerous studies in the experimental literature use such structural specifications, as they permit errors but do not increase the number of parameters to be estimated (for example, Harrison and Rutström 2009, Harrison et al. 2010). Determining the extent of noise or randomness in experimental decision-making is not a focus of this chapter, and is explored in Chapters 1 and 3. See Appendix A and B in Chapter 1 for more details on this model.

Thus, given the observed choice y_a^d of each individual a in decision problem d, the log-likelihood of the observed responses – conditional on the EUT-CRRA specification described in Equations (2) and (3) representing the true decision-making process of subjects[13] – is:

$$\ell^{EUT} = \ln L^{EUT}(\delta, r; \boldsymbol{y}) = \sum_{a=1}^{N_d} \ln(\nabla EU_{y_a^d}^d) \quad (6)$$

where N_d is the number of participants in decision problem d and $\boldsymbol{y}$ is the vector of observed choices in the decision problem.[14]

This structural estimation methodology is also used by Harrison et al. (2007) and Heinemann (2008) to analyze asset integration. However, these studies do not possess information on subjects' outside wealth W, and hence maximize the log-likelihood function ℓ^{EUT} with respect to integrated wealth δW (as opposed to δ) and r – this is equivalent to including a constant term in the argument of the utility function in addition to x, and then estimating this constant term jointly with the parameter r. However, the ERHS dataset contains information on subjects' outside wealth – using this wealth data, the log-likelihood function ℓ^{EUT} can be maximized with respect to the parameters δ and r. Therefore, the information on household wealth contained in the ERHS dataset enables the separate identification of the degree of asset integration δ (rather than the just δW); building on the work of Harrison et al. (2007) and Heinemann (2008), I obtain maximum likelihood estimates of δ and r, thus enabling a more accurate estimation of the degree of asset integration. Andersen et al. (2006a) and Harrison et al. (2011) conduct similar estimations using experimental data from the United States and Denmark, respectively. As mentioned earlier, estimates of δ close to 0 provide evidence of narrow framing, as opposed to asset integration.

[13] Note that here I follow the representative agent assumption that the choices of all subjects are generated by the same model of decision-making under risk (with the same preference parameter values).

[14] Note that here I assume that all the choices in the decision problem are independent, and generated by a single model of decision-making under risk, following the strategy of Harrison et al. (2007) and Harrison and Rutström (2008b).

In addition to these estimations, I also treat δW as a constant term – as done by Harrison et al. (2007) and Heinemann (2008) – and conduct maximum likelihood estimations involving the optimization of the log-likelihood function with respect to δW and r. This would shed light on the robustness of the results obtained from the specifications involving the estimation of δ, and also enable a direct comparison of the results with those of the above-mentioned studies. Values of the constant term δW close to 0 provide empirical support for the hypothesis that participants maximize utility defined only over lottery prizes (narrow framing); however, this specification does not provide information on the degree of asset integration (since it does not involve the separate estimation of δ). Both these estimations let the "data decide" what arguments of the utility function best describe observed behaviour, without imposing the assumption of narrow framing or full asset integration *a priori*, as done by most studies of decision-making under risk (Andersen et al. 2006a).

2.2 Determinants of risk aversion in an EUT framework

In Chapter 1, and so far in Chapter 2, I have made the traditional representative agent assumption that each subject in the experiment has the same preferences – each participant's choices under risk are assumed to be generated (with error) by the same latent decision-making process (that is, the same decision model with the same functional parameter values); this assumption has also been incorporated by Andersen et al. (2006a) and Harrison et al. (2011) in the estimation of risk preferences.[15] However, Guiso and Paiella (2008) and Tanaka et al. (2010) find that there is fundamental heterogeneity in risk preferences, thus casting doubt on the validity of representative agent models for describing choices involving risk; in addition, Harrison and Rutström (2008b) stress the importance of allowing for heterogeneity in risk preferences within a given decision theory. Therefore, following the strategy of Harrison et al. (2007) and Harrison and

[15] Hey (1995) notes that the representative agent in these models can be regarded as the average subject involved in the experiment, and thus parameter estimates of the decision models can be interpreted as the parameter values for the average subject.

Rutström (2008b), I allow each preference parameter of the decision model to be a linear function of various demographic and socioeconomic characteristics of the subject and his household, in order to account for the heterogeneity in risk attitudes and analyze the impact of these variables on risk preferences.

Therefore, I assume that subjects evaluate decisions using the EUT-CRRA specification described in Equations (1) and (2); x is now assumed to be the only argument of the utility function, unlike in the asset integration specification of the previous section. Further, I no longer assume that each subject is associated with the same value of r, but instead consider r to be a linear function of various participant characteristics. Therefore, for each individual a:

$$U_a(x) = \begin{cases} \dfrac{x^{1-r_a}}{1-r_a} & \text{if } r_a \neq 1 \\ \ln(x) & \text{if } r_a = 1 \end{cases} \tag{7}$$

Using this utility function, expected utility from lottery i for individual a (EU_{ai}) is computed as the sum of the utility for individual a (U_a) from each outcome in the lottery weighted by the associated objective probability, following Equation (2).

For each participant a, $r_a = r(\boldsymbol{X})$, where $\boldsymbol{X}$ is a (column) vector of individual and household characteristics of the participant that are *a priori* believed to affect risk preferences; thus, the risk aversion of each participant is a linear function of $\boldsymbol{X}$. The two explanatory variables for risk aversion that I focus on in this analysis are wealth (W) and background risk ($Brisk$). Background risk is a measure of risk that is statistically independent of subjects' earnings – and thus risk (labelled "foreground" risk) – in the experimental decision problems.[16] Therefore, r_a can

[16] As highlighted in Section 1, the *a priori* expectation of the sign of impact of wealth and background risk on risk preferences is ambiguous. Given EUT-CRRA, a negative estimated impact of background risk on risk aversion would represent evidence consistent with the theoretical predictions of Quiggin (2003), while a positive estimated impact would be in line with Gollier and Pratt's (1996) theory of risk vulnerability.

be described by:

$$r_a = \boldsymbol{\beta X_a} + \varepsilon_a \tag{8}$$

$$r_a = \beta_0 + \beta_1 W_a + \beta_2 Brisk_a + \boldsymbol{\beta_3 X_{2a}} + \varepsilon_a \tag{9}$$

where ε_a is the error term, and $\boldsymbol{X_{2a}}$ is the (column) vector consisting of the characteristics of participant a excluding wealth and background risk. Thus, $\boldsymbol{X_{2a}}$ includes variables such as gender, age, occupation and education, which are also expected to affect risk preferences – these variables are described in detail in Section 3. Using this specification, the estimation of r_a entails the estimation of the coefficients β_0 (a constant), β_1, β_2 and $\boldsymbol{\beta_3}$ (which is a row vector of coefficients that correspond to the characteristics comprising $\boldsymbol{X_{2a}}$); following common practice, these coefficients are assumed to be the same for all participants (Harrison and Rutström 2008b).

Assuming an EUT-CRRA specification (given by Equations (2) and (7)), the choice of each individual a in an experimental decision problem implies a choice of, and corresponds to, a particular range of the coefficient of relative risk aversion r_a. This is the range of coefficients for which that particular choice would be optimal for an individual with EUT-CRRA preferences. Therefore, the dependent variable r_a itself is not observed, and only the interval that contains it can be inferred from subjects' choices (Bettin and Lucchetti 2010). That is, each experimental choice only enables the observation of the bounds that contain r_a – m_a and M_a – where

$$m_a < r_a \leq M_a \tag{10}$$

and the interval may be right- or left-unbounded.

In order to estimate the (row) vector $\boldsymbol{\beta}$ of coefficients listed above, I use an interval regression model, as done by Wik et al. (2004), Lusk and Coble (2008) and Tanaka and Munro (2012) in similar contexts. With the assumption of EUT-CRRA preferences, the interval regression model utilizes information on the bounds of relative risk aversion implied by observed choices in the decision problems (Harrison and Rutström 2008b). In reference to Equation (8), the model assumes that the error term is independently and identically normally distributed with mean 0 and standard deviation σ; that is, $\varepsilon_a \sim \mathcal{N}(0, \sigma^2)$ (Wooldridge 2002).[17] With this distributional assumption, the coefficient vector $\boldsymbol{\beta}$ can be estimated using maximum likelihood techniques. As noted by Bettin and Lucchetti (2010), the log-likelihood function for the choice y_a of subject a, given the implied bounds of relative risk aversion (m_a and M_a) and the assumption of normality of the error, is given by:

$$\ell_a = \ln L_a(\boldsymbol{\beta}, \sigma; y_a, \boldsymbol{X_a}) = \ln \text{Prob}(m_a < r_a \leq M_a | \boldsymbol{X_a}) \tag{11}$$

$$= \ln \left[\Phi \left(\frac{M_a - \boldsymbol{\beta X_a}}{\sigma} \right) - \Phi \left(\frac{m_a - \boldsymbol{\beta X_a}}{\sigma} \right) \right] \tag{12}$$

where $\boldsymbol{X_a}$ is the vector of values of the characteristics of individual a, and m_a and M_a are the bounds of the coefficient of relative risk aversion that correspond to choice y_a (assuming EUT-CRRA preferences). The individual choice log-likelihood function is then summed over the choices in the decision problem(s) considered in the estimation, yielding the total log-likelihood function. The total log-likelihood function can then be maximized with respect to $\boldsymbol{\beta}$ and σ to obtain the maximum likelihood estimates of the parameters. Using the vector of coefficient estimates $\hat{\boldsymbol{\beta}}$, the estimate of subject a's coefficient of relative risk aversion is $\hat{r}_a = \hat{\boldsymbol{\beta}} \boldsymbol{X_a}$.[18]

However, the model outlined above does not account for the possible endogeneity associated with the explanatory variables wealth and background risk (highlighted in Section 1). In order to account for this endogeneity and consistently estimate $\boldsymbol{\beta}$, I utilize two extensions of the basic interval regression model that incorporate instrumental variables – these are described by Bettin and Lucchetti (2010), and build on the work of Manski and Tamer (2002) and Hong

[17] For $\boldsymbol{\varepsilon}$, the vector comprising the error terms for each individual, $\boldsymbol{\varepsilon} \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{I})$ (Wooldridge 2002).

[18] In the interval regression model, the coefficient estimates of the characteristics included in $\boldsymbol{X}$ are interpreted as estimates of the impact of these characteristics on risk aversion r (Harrison and Rutström 2008b).

and Tamer (2003). Following closely the notation of Bettin and Lucchetti (2010) and letting $\boldsymbol{Y_a}$ represent the (column) vector comprising the two endogenous variables for individual a – that is, W_a and $Brisk_a$ – the model can be expressed as:

$$r_a = \boldsymbol{\alpha Y_a} + \gamma \boldsymbol{X_{2a}} + \varepsilon_a \tag{13}$$

$$\boldsymbol{Y_a} = \boldsymbol{\Pi_1 X_{2a}} + \boldsymbol{\Pi_2 Z_a} + \boldsymbol{u_a} = \boldsymbol{\Pi S_a} + \boldsymbol{u_a} \tag{14}$$

and

$$\begin{bmatrix} \varepsilon_a \\ \boldsymbol{u_a} \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \sigma_\varepsilon^2 & \boldsymbol{\theta}' \\ \boldsymbol{\theta} & \boldsymbol{\Sigma} \end{bmatrix}\right)$$

where $\boldsymbol{Z_a}$ is the (column) vector of instrumental variables for wealth and background risk[19], $\boldsymbol{\Sigma}$ is the covariance matrix and $\boldsymbol{u_a}$ is the (column) vector containing the two error terms of the wealth and background risk specifications (in Equation (14)). $\boldsymbol{\theta}$ is the covariance between ε_a and $\boldsymbol{u_a}$, which may be non-zero – if this is the case, $\boldsymbol{Y_a}$ is indeed endogenous and the basic interval regression model outlined earlier does not provide consistent estimates of $\boldsymbol{\alpha}$ and γ (Bettin and Lucchetti 2010).

Along the lines of the consistent estimator for endogenous probit models described by Rivers and Vuong (1988), a two-step estimator to account for endogeneity in interval regression models is proposed by Bettin and Lucchetti (2010). The first stage involves running an Ordinary Least Squares (OLS) regression of $\boldsymbol{Y_a}$ on $\boldsymbol{S_a}$, in order to estimate $\boldsymbol{\Pi}$ and $\boldsymbol{\Sigma}$ and obtain the predicted residual $\boldsymbol{\hat{u}_a}$. In the second stage, $\boldsymbol{\alpha}$ and γ are estimated by running a basic interval regression using the bounds m_a and M_a (implied by observed choice), with $\boldsymbol{Y_a}$, $\boldsymbol{X_{2a}}$ and $\boldsymbol{\hat{u}_a}$ as the explanatory variables. Bettin and Lucchetti (2010) note that this two-step estimator is consistent.

Consistent parameter estimates can also be obtained using a Limited Information Maximum Likelihood (LIML) estimator, as described by Bettin and Lucchetti

[19] These instrumental variables are described in detail in Section 3.1.

(2010). They note that the log-likelihood function can be split into a conditional component and marginal component:

$$\ell_a(\boldsymbol{\Psi}) = \ell_a^C + \ell_a^M \tag{15}$$

where $\boldsymbol{\Psi}$ is a vector containing all the parameters to be estimated. Given the bounds of risk aversion – m_a and M_a – that correspond to a particular choice y_a of individual a (since r_a is imperfectly observed), the conditional component is given by:

$$\ell_a^C(\boldsymbol{\Psi^*}) = \ln \text{Prob}(m_a < r_a \leq M_a | \boldsymbol{X_a}, \boldsymbol{u_a}) \tag{16}$$

$$= \ln \left[\Phi \left(\frac{M_a - \hat{r}_a}{\sigma} \right) - \Phi \left(\frac{m_a - \hat{r}_a}{\sigma} \right) \right] \tag{17}$$

where

$$\hat{r}_a = \boldsymbol{\alpha Y_a} + \gamma \boldsymbol{X_{2a}} + \boldsymbol{\Omega u_a} \tag{18}$$

Bettin and Lucchetti (2010) note that the marginal component in this model is an ordinary multivariate Gaussian log-likelihood:

$$\ell_a^M = \ln f(\boldsymbol{u_a}, \boldsymbol{\Psi^*}) = -\frac{1}{2} \left[2 \ln(2\pi) + \ln |\boldsymbol{\Sigma}| + (\boldsymbol{Y_a} - \boldsymbol{\Pi S_a})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{Y_a} - \boldsymbol{\Pi S_a}) \right] \tag{19}$$

The individual choice log-likelihood function $\ell_a(\boldsymbol{\Psi})$ is summed over all choices, and the resultant total log-likelihood function can then be maximized (numerically) with respect $\boldsymbol{\Psi}$ (using the analytical score) to obtain LIML estimates of the parameters comprising the vector $\boldsymbol{\Psi}$.[20] Bettin and Lucchetti (2010) note that the LIML estimator defined above is also consistent, like the two-step estimator, and it coincides with the two-step estimator in the just-identified case. They also note that the relative efficiency of these two estimators is comparable.[21]

[20] Details on the computation of the analytical score are provided in Appendix A of Bettin and Lucchetti (2010).

[21] Since both the two-step and LIML models use data on the intervals of r in addition to instrumental variables, I refer to them jointly as IV interval regression models.

Another commonly-used method to account for endogeneity when the dependent variable is observed in the form of intervals is to approximate the variable using the midpoint of the bounds, and then run the standard Two-Stage Least Squares procedure using this midpoint data and instrumental variables (Bettin and Lucchetti 2010). However, using Monte Carlo simulations, Bettin and Lucchetti (2010) find that the application of Two-Stage Least Squares to the midpoint dependent variable yields inconsistent estimates, as well as substantial errors in small samples. On the other hand, they find that the two-step and LIML estimators are consistent, and also produce accurate estimates for relatively small samples[22]; however, despite their advantages, such IV interval regression models have rarely been used in the applied literature.

In Chapter 1, I found evidence that EUT with CRRA utility provides a good description of the choices in the Binswanger-style lottery of the Ethiopian experiment (decision problem B), while decision-making in the insurance decision problems (decision problems T_{IM}, T_{GM}, T_{IX} and T_{GX}) is better described by non-EUT models, in particular a two-part power utility function with loss aversion (but no probability weighting). Therefore, I use the models described in Sections 2.1 and 2.2 – which assume EUT-CRRA as the latent decision-making process of individuals – to analyze choices from Binswanger-style lotteries in both the Ethiopian experiment and the ERHS. However, the interval regression model and IV interval regression extensions are only suitable for analyzing choices using single-parameter decision models, such as EUT-CRRA, and therefore cannot be used for the estimation of non-EUT models which incorporate multiple parameters (Harrison and Rutström 2008b, Lusk and Coble 2008). Thus, to study the determinants of risk preferences exhibited in the insurance decision problems of the Ethiopian experiment, in which choices are likely to be better characterized by a non-EUT specification, I use a two-step structural maximum likelihood estimator, as described in the following section.

[22] However, the standard errors are very similar in all three cases.

2.3 Determinants of risk preferences in a non-EUT framework

To jointly study the determinants of risk attitudes exhibited in all four insurance decision problems of the Ethiopian experiment, I assume that the choices in these problems are generated by a two-part power utility specification with loss aversion (but no probability weighting), which in Chapter 1 was found to be the most likely candidate for representing the true decision-making process generating choices in these decision problems. In particular, in Chapter 1 I found that the two-part power utility specification and RDU with CRRA utility may both be suitable for providing an appropriate description of choices in the insurance problems of the Ethiopian experiment. However, using Vuong (1989) and Clarke (2003) tests for non-nested models and data from the insurance problems, the RDU-CRRA specification is rejected in favour of the two-part power utility specification (with no probability weighting) at the 1% significance level, implying that, according to these tests, the two-part power utility specification is closer to the true data-generating process generating choices in the insurance problems; in other words, the two-part power utility specification better describes the true decision-making process of individuals in the insurance problems and thus fits the data better.[23]

The two-part power utility function for individual a, which is defined separately over gains and losses and is described in detail in Chapter 1, has the following form:

$$U_a(x,\chi) = \begin{cases} (x-\chi)^{\alpha_a} & \text{if } (x-\chi) \geq 0 \\ -\lambda_a(-(x-\chi))^{\alpha_a} & \text{if } (x-\chi) < 0 \end{cases} \tag{20}$$

Using this utility function, gain-loss utility from lottery i for individual a (GLU_{ai}) is computed as the sum of the utility for individual a (U_a) from each outcome in the lottery weighted by the associated objective probability, following

[23] It is important to note that the Vuong (1989) and Clarke (2003) tests are tests of specific parametric forms of the decision theories, rather than general tests of the theories themselves.

Equation (2).

Thus, utility is no longer defined over the final prize amounts (x) in the lotteries, but is defined separately over gains and losses relative to a reference point χ.[24] Further, α is the parameter capturing the curvature of the utility function (referred to as the risk aversion parameter by Harrison and Rutström 2009). Values of $\alpha < 1$ indicate a utility (or value) function which is concave above the reference point and convex below it. That is, $\alpha < 1$ reflects risk aversion for prize amounts greater than the reference point (gains), and risk seeking behaviour for prize amounts less than the reference point (losses).

λ is the loss aversion parameter. $\lambda > 1$ indicates loss aversion, and the magnitude of λ determines the extent to which losses loom larger than gains in decision-makers' valuations of risky lotteries, with higher values of λ indicating greater loss aversion. Further, loss aversion generates a kink at the reference point which makes the utility function steeper for losses than for gains around the reference point (see Chapter 1 for more details). When $\lambda = 1$, there is no loss aversion, and gains and losses are evaluated equivalently. This form of utility function was used by Tversky and Kahneman (1992) in their development of cumulative prospect theory.

As done in the previous section, I consider the parameters of the utility function – α and λ – to be linear functions of participant characteristics. Thus, for each participant a, $\alpha_a = \alpha(\boldsymbol{X})$ and $\lambda_a = \lambda(\boldsymbol{X})$, where $\boldsymbol{X}$ is the same (column) vector of individual and household characteristics considered in the previous section. Following Harrison et al. (2007) and Harrison and Rutström (2008b), α_a and λ_a

[24] Following Harrison and Rutström (2008b), and the work in Chapter 1, I assume a common reference point χ for all subjects.

can therefore be described by:

$$\alpha_a = \boldsymbol{\beta X_a} \tag{21}$$

$$\alpha_a = \beta_0 + \beta_1 W_a + \beta_2 Brisk_a + \boldsymbol{\beta_3 X_{2a}} \tag{22}$$

$$\lambda_a = \boldsymbol{\tau X_a} \tag{23}$$

$$\lambda_a = \tau_0 + \tau_1 W_a + \tau_2 Brisk_a + \boldsymbol{\tau_3 X_{2a}} \tag{24}$$

where $\boldsymbol{X_{2a}}$ is the (column) vector of characteristics of participant a excluding wealth (W_a) and background risk ($Brisk_a$).

Further, in Chapter 1, I obtained an estimate of 17.076 birr for the reference point in the two-part power utility function, using the structural maximum likelihood estimation technique to analyze choices in all four insurance decision problems. Therefore, in the two-part power utility estimations in this chapter, I assume a reference point of 17.076 birr for the insurance decision problems. Additionally, while the original formulation of cumulative prospect theory used a two-part power utility function of this form in combination with cumulative probability weighting, I do not allow for the weighting of probabilities in these estimations, because in Chapter 1 I did not find evidence of probability weighting – when cumulative probability weighting functions were combined with the two-part power utility function – in the decision-making of subjects in the insurance decision problems. Along similar lines, Tanaka et al. (2010) only analyze the determinants of λ and α in their prospect theory specifications (and assume an exogenously-given reference point of 0).

Using this specification, the estimation of α_a and λ_a entails the estimation of coefficients $\boldsymbol{\beta}$ and $\boldsymbol{\tau}$. To estimate these coefficients, I use the structural maximum likelihood estimation procedure, which is also used extensively in Chapter 1 and is summarized below. First, the gain-loss utility GLU^d_{ai} for individual a from each potential lottery choice $i \in \{1, \ldots, 6\}$ in insurance decision problem $d \in \{T_{IM}, T_{GM}, T_{IX}, T_{GX}\}$ is calculated according to Equation (2) (objective probability weighting), assuming that utility is defined by the two-part power utility function in Equation (20) (with a reference point χ of 17.076 birr). The latent index ∇GLU_{ai}^{d} for a particular choice of i individual a is then calculated as follows:

$$glu_{ai}^{d} = \exp(GLU_{ai}^{d}) \tag{25}$$

$$\nabla GLU_{ai}^{d} = \frac{glu_{ai}^{d}}{\sum\limits_{b=1}^{6} glu_{ab}^{d}} \tag{26}$$

The latent index ∇GLU_{ai}^{d}, based on latent preferences, is in the form of a probability between 0 and 1, and can be directly linked to observed choices.[25] Thus, given the observed choice y_{a}^{d} of each individual a in a particular insurance decision problem d, the log-likelihood of the observed responses in d, conditional on the two-part power utility specification described in Equations (2) and (20) representing the true decision-making process of subjects, is:

$$\ell_{d}^{GLU} = \ln L_{d}^{GLU}(\alpha, \lambda; \boldsymbol{y}, \boldsymbol{X}) = \sum_{a=1}^{N_d} \ln(\nabla GLU_{ay_{a}^{d}}^{d}) \tag{27}$$

where N_d is the number of participants in insurance decision problem d, and $\boldsymbol{y}$ is the vector of observed choices in d. Since I aim to jointly analyze the choices in all four insurance decision problems (as done in Chapter 1), the total log-likelihood for all choices in the four insurance decision problems is obtained by summing the log-likelihood functions for each insurance decision problem d, as follows:

$$\ell^{GLU} = \ln L^{GLU}(\alpha, \lambda; \boldsymbol{y}, \boldsymbol{X}) = \sum_{d} \sum_{a=1}^{N_d} \ln(\nabla GLU_{ay_{a}^{d}}^{d}) \tag{28}$$

where $\boldsymbol{y}$ is now the vector consisting of the decisions in all four insurance decision problems. This log-likelihood function can then be optimized numerically with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\tau}$ in order to obtain the structural maximum likelihood estimates of the coefficients corresponding to the elements of $\boldsymbol{X}$, thus providing

[25] This once again represents a model of "stochastic choice with deterministic preferences" with the extent of errors in decision-making fixed at a certain level, similar to the model described in Equation (5).

estimates of α_a and λ_a.[26] Therefore, this specification enables the analysis of the determinants of the loss aversion parameter λ (in addition to the determinants of the curvature parameter α) – this is an important aim of this chapter. As noted by Tanaka et al. (2010), there is a distinct lack of experimental research on the determinants of risk preferences (and loss aversion in particular) using non-EUT decision models.

The estimates from this model, however, may be inconsistent due to the endogeneity associated with wealth and background risk.[27] Therefore, in order to account for this endogeneity and obtain consistent coefficient estimates, I use a two-step estimator along the lines of the two-step interval regression estimator described in Section 2.2. Expressing the model with endogeneity as done in Equations (13) and (14), the first step of this two-step structural maximum likelihood estimator is identical to that of the two-step interval regression estimator. Using the notation employed in Equations (13) and (14), the first step involves running an OLS regression of $\boldsymbol{Y_a}$ on $\boldsymbol{S_a}$, in order to estimate $\boldsymbol{\Pi}$ and $\boldsymbol{\Sigma}$ and obtain the predicted residual $\boldsymbol{\hat{u}_a}$. However, in the second stage, instead of using a basic interval regression model, I utilize the structural maximum likelihood model described above, with $\boldsymbol{Y_a}$, $\boldsymbol{X_{2a}}$ and $\boldsymbol{\hat{u}_a}$ as the explanatory variables for α_a and λ_a. That is, I allow α_a and λ_a to be linear functions of $\boldsymbol{Y_a}$, $\boldsymbol{X_{2a}}$ and $\boldsymbol{\hat{u}_a}$, and use the structural maximum likelihood model to estimate the coefficients corresponding to these explanatory variables. Like the two-step interval regression estimator, the two-step structural maximum likelihood

[26] As noted by Harrison and Rutström (2008b), this approach assumes that all of the heterogeneity between participants is captured by the individual and household characteristics included in the specification; this may be a limitation, as even with the inclusion of an extensive set of socioeconomic and demographic variables, there could be a concern that the risk preferences are affected by unobserved individual heterogeneity. Harrison and Rutström (2008b) also note that while one possible solution to this problem is to conduct a within-subject analysis, which implicitly controls for unobserved heterogeneity, this requires a very large sample of choices for each participant (for example, the choices of each subject in a number of Multiple Price List decision problems).

[27] Indeed, Liu (2008), using a similar two-part power utility function to describe the experimental choices of Chinese farmers, finds that both the loss aversion parameter (λ) and the curvature parameter (α) influence technology adoption decisions, which in turn are likely to affect wealth and background risk. This highlights the importance of accounting for the possible endogeneity associated with wealth and background risk (which could arise due to reverse causality) when estimating their impact on loss aversion and the curvature of the utility function.

estimator is also expected to be consistent.

3 Data and Descriptive Statistics

The dataset used in this study is a combination of the experimental data from Ethiopia used in Chapter 1 and data from the Ethiopian Rural Household Survey (ERHS). The experiment, which is described in detail in Chapter 1, involved 378 subjects from seven sites of the ERHS, spanning three regions of the country.[28] The ERHS dataset is a longitudinal household dataset covering approximately 1,350 households located in 15 Ethiopian villages, and contains data from a survey conducted over seven rounds from 1994 to 2009. There were six original villages chosen for the ERHS in 1989 because of the drought that affected Ethiopia in 1984-85; the rest were added in 1994 to achieve broad diversity of rural Ethiopia. The data collection for the ERHS was coordinated by the Economics Department at Addis Ababa University in collaboration with the Centre for the Study of African Economies at Oxford University and the International Food Policy Research Institute. For the purpose of this chapter, I use data mostly from the latest round (round 7) of the survey, which was conducted between April and August 2009, since the timing closely matches that of the Ethiopian experiment, which was conducted during November and December 2009. However, certain panel aspects of the data are also utilized, and will be described later in this section.

The ERHS dataset contains detailed information on various socioeconomic and demographic characteristics of households in the sample, such as asset ownership, consumption, occupation and household size.[29] Even though the

[28] The seven ERHS sites chosen for the experiment were Sirbana Godeti, Korodegaga, Indibir, Milki, Komargefia, Karafino and Bokafia. These sites are close in distance to the capital of Addis Ababa (and span three regions of the country), and within these sites, all households who participated in the ERHS were invited to participate in the experiment. Approximately 96% of invited households attended the experiment.

[29] For more details on the survey and sampling procedures, see Dercon and Krishnan (1998) and Dercon and Hoddinott (2004).

Table 1. *Binswanger-style decision problem in round 7 of the ERHS*

Choice A:	Receive 2.5 birr regardless of whether the outcome is a lion or a crown
Choice B:	Receive 2 birr if the outcome is a lion, 4 birr if the outcome is a crown
Choice C:	Receive 1.5 birr if the outcome is a lion, 5.5 birr if the outcome is a crown
Choice D:	Receive 1 birr if the outcome is a lion, 7 birr if the outcome is a crown
Choice E:	Receive 0 birr if the outcome is a lion, 10 birr if the outcome is a crown

ERHS households are not nationally representative, the survey covered the primary agro-climatic zones of the country (Hill et al. 2011), and has been used extensively in published studies on various aspects of the rural Ethiopian economy (for example, Dercon and Krishnan 2000, Dercon 2004, Fafchamps and Quisumbing 2005).

The Ethiopian experiment involved a Binswanger-style lottery (benchmark decision problem B) and four insurance decision problems (T_{IM}, T_{IX}, T_{GM} and T_{GX}) – these decision problems are described in detail in Chapter 1. In addition to subjects' choices in these decision problems, the experimental dataset contains information on schooling, quantitative literacy, occupation, age and various other demographic characteristics of the participants.

Further, all the experimental subjects were from households included in the ERHS. Therefore, the data from the experimental sessions was matched to the data from the ERHS for each experimental subject using unique household identification numbers, thus enabling the creation of a combined dataset with both data from ERHS and the experiment for each subject involved in the experiment. In addition, the latest round of the ERHS (carried out in 2009) included a module on risk preferences, in which surveyed individuals were presented a gain frame Binswanger-style lottery with real (monetary) payoffs. In particular, they were asked to choose which of the five gambles presented in Table 1 they most preferred, after which their preferred gamble was played for real money and they were paid based on the outcome of a fair coin flipped by the enumerator.

If the decision-making process of subjects is assumed to be described by an EUT-

Table 2. *Risk aversion ranges in the ERHS Binswanger-style decision problem*

	Choice	Net Payoff (Ethiopian birr)		Expected Payoff	Risk Aversion Range (CRRA)*
Probability:		$1/2$	$1/2$		
	A	2.5	2.5	2.5	$(3.26, +\infty)$
	B	2	4	3	$(1.2, 3.26)$
	C	1.5	5.5	3.5	$(0.68, 1.2)$
	D	1	7	4	$(0.33, 0.68)$
	E	0	10	5	$(-\infty, 0.33)$

* Risk aversion range denotes range of coefficients for which choice would be optimal for a subject with CRRA preferences (Equation (1)) over earnings from the ERHS Binswanger-style lottery.

CRRA specification (as described in Equations (1) and (2)), then each choice in this Binswanger-style lottery implies a choice of, and corresponds to, a particular range of the coefficient of relative risk aversion r. These ranges of r are provided in Table 2, which also shows that the lottery options increase in both riskiness (variance) and expected payoffs from A to E. Similar tables for the decision problems in the Ethiopian experiment are provided in Chapter 1.

Thus, the dataset used in this study contains the choices in this Binswanger-style lottery of nearly 1,350 ERHS participants and the choices of the 378 experiment participants in the five decision problems of the Ethiopian experiment[30], as well as detailed individual- and household-level survey data on all these participants. This combined experimental and survey dataset provides fertile ground to investigate the influence of an extensive set of demographic and socioeconomic factors – including wealth and background risk – on decision-making and risk-taking in the various decision problems. Such combined datasets have rapidly gained popularity in the empirical literature and have been used, for example, to test the effect on risk attitudes of wealth (Mosley and Verschoor 2005, Yesuf and Bluffstone 2009, Tanaka et al. 2010) and agro-climatic conditions (Tanaka and Munro 2012), and to analyze the asset integration hypothesis (Schechter 2007). Additionally, they have also been used extensively to study the impact of

[30] In the Ethiopian experiment, all 378 participants made choices in B, while 136 made choices in T_{IM}, 258 made choices in T_{IX}, 120 made choices in T_{GM} and 242 made choices in T_{GX}.

experimentally-measured risk preferences on various important non-experimental economic decisions – such as technology adoption and risk pooling decisions – in developing country settings (for example, Barr and Packard 2005, Engle-Warnick et al. 2011, Attanasio et al. 2012).

In order to analyze the determinants of risk preferences using the models described in Sections 2.2 and 2.3, I use a number of household-level and respondent-level (or individual-level) characteristics as explanatory variables to describe risk preferences – these characteristics are denoted by the vector $\boldsymbol{X}$ in the model descriptions. The summary statistics for these explanatory variables – that is, the variables whose effect on risk preferences I estimate – are provided in Table 3; the summary statistics are separately presented for the experiment sample and complete ERHS round 7 sample.[31] The household-level variables include wealth, background risk and household size, while the respondent-level variables include age, gender, quantitative literacy, formal education, occupation and whether the respondent is a household head or not.

The two explanatory variables I focus on in this analysis are wealth and background risk. I use two alternative measures of household wealth – tropical livestock units and total land owned. Tropical livestock units are standardized units of different types of livestock, and they are used as a measure of total livestock ownership in numerous studies set in the context of developing economies (for example, Dercon 2004, Barrett and McPeak 2006, Lybbert et al. 2010).[32,33] Dercon (2004), in a study of growth and shocks in rural Ethiopia, finds that livestock typically accounts for over 90% of the value of household

[31] Note that sample selection bias is unlikely, since experiment participants were chosen from Ethiopian Rural Household Survey (ERHS) households purely on the basis of location. Experimental subjects were chosen from seven sites of the ERHS that are close in distance to the capital of Addis Ababa (and span three regions of the country). Within these sites, all households who participated in the ERHS were invited to participate in the experiment; approximately 96% of invited households attended the experiment.

[32] I use the terms tropical livestock units and total livestock units interchangeably in this chapter.

[33] Tropical livestock units provide a single figure that expresses the total amount of livestock owned, allowing different animals to be described in relation to a common unit. For the purposes of the ERHS, it is calculated using the following conversions: oxen=1, cows=0.70, bulls=0.75, horses=0.50, goats=0.10, sheep=0.10 and other similar values (Dercon 2004).

Table 3. *Summary statistics*

	Experimental Sample			Complete ERHS round 7 sample		
Variable	No. of Observations	Mean	Standard Deviation	No. of Observations	Mean	Standard Deviation
Household-level Variables						
Tropical livestock units	372	10.54	2.900	1348	9.300	2.070
Total land owned[a]	372	3.033	4.844	1348	1.912	2.909
Std. dev. of consumption[b]	370	528.8	335.4	1348	406.8	303
Household size	372	5.551	2.341	1348	5.700	2.562
Individual-level Variables						
Age (years)	378	45.14	15.94	1271	51.67	15.24
Gender (1 if female, 0 if male)	378	0.325	0.469	1271	0.292	0.455
Formal schooling obtained (years)	377	4.149	3.795	1270	1.739	3.039
Quantitative literacy	378	0.576	0.195	1342	0.597	0.236
Household head (1 if household head, 0 if not)	378	0.696	0.461	1271	0.947	0.224
Farmer (1 if farmer, 0 if not)	378	0.664	0.473	1269	0.737	0.441

[a] Measured in hectares
[b] Measured in 1994 Ethiopian birr

assets and is the most marketable asset in this region – therefore, in rural Ethiopia, as in many of the poorest rural regions of the world, livestock ownership is an appropriate and important measure of household wealth, while total livestock units provide a suitable and comparable description of livestock ownership. The average livestock ownership in the round 7 ERHS sample is 9.3 total livestock units, which is slightly less than that in the experimental sample (10.5 total livestock units).

Following Porter (2011) and Hill et al. (2011), I also use the total land owned by the household (measured in hectares), including agricultural and non-agricultural land, as an alternate measure of wealth – with agricultural production as the primary occupation in the region, land is certainly a crucial asset for households in Ethiopia (Ayalew et al. 2005). The average household landholding in the ERHS sample is 1.9 hectares – this is larger than the national average of 0.9 hectares per household (Hill et al. 2011), but is substantially less than the average in the experimental sample (3 hectares).[34,35] Both land and livestock ownership are also used to test the hypothesis of asset integration using the model described in Section 2.1.

Exploiting the panel aspect of the ERHS dataset, I use the standard deviation of household consumption over all seven rounds of the ERHS (conducted between 1994-2009) as a measure of the level of background risk to consumption faced by the participant's household in recent past. Numerous studies use measures based on the inter-temporal standard deviation of consumption or income as measures of risk and shocks affecting rural households (for example, Jalan and Ravallion 1998, Kamanou and Morduch 2004). In this case, the standard deviation of consumption is used rather than that of income because the ERHS income data is far less reliable than the consumption data (Porter 2011). In

[34] Additionally, it is worth noting that all households own at least some land (as well as some livestock).

[35] The difference between the average household landholding for the ERHS and experimental samples is likely because experimental subjects were chosen from seven sites of the ERHS that are close in distance to the capital of Addis Ababa, and these villages are expected to be wealthier, on average, than those further away from the capital city (Dercon and Hoddinott 2004).

addition, I want to measure the final (residual) level of risk faced by the household – that which could have a significant impact on household welfare – after it has used all available insurance and consumption smoothing measures (including possibly-inefficient asset investment and divestment), which would not be captured by the standard deviation of income.

In the context of decisions in the ERHS or experiment lotteries, this measure can be considered as background risk, as it is uncorrelated with the earnings and foreground risk in the lotteries and is difficult to insure against (since I use consumption, which provides a measure of welfare after the use of available insurance mechanisms and risk-coping strategies to smooth income).[36] For a particular decision problem in the experiment or ERHS, all subjects face the same risky outcomes and are given the same choices, but differ significantly in their exposure to background risk (this is reflected by the relatively large inter-household variation in the standard deviation of consumption measure, as indicated by Table 3).[37] Consumption is measured as the total monthly household consumption in 1994 Ethiopian birr. It includes the consumption of food, purchased food and non-investment non-food items (that is, it excludes expenditure on durables, health and education) – as noted by Porter (2011), this measure has been used in various other studies of consumption and poverty conducted using the ERHS dataset. Table 3 shows that the average inter-temporal standard deviation of consumption of around 400 birr represents a significant fraction of the average consumption measured in round 7 (approximately 1100 birr). This indicates that the subjects considered in this analysis face considerable risk to consumption.

[36] Further, this measure of background risk captures the risk participants' face in their day-to-day lives outside the experiment.

[37] Guiso and Paiella (2008) and Hill et al. (2011) also use variables based on the inter-temporal deviation of household consumption to measure the background risk faced by individuals in their samples (with the latter using data from the rounds of the ERHS from 1994-2009). However, it is important to note that while the inter-temporal deviation in household consumption does capture background risk faced by the participant, all of the deviation in consumption is not due to risk – for example, some of the inter-temporal variation in household consumption could be caused by life cycle effects or household decisions regarding the purchase of durables.

Table 3 also provides the summary statistics for other demographic variables that are expected to affect risk attitudes. These variables comprise the vector $\boldsymbol{X_{2a}}$ referred to in Sections 2.2 and 2.3. Such demographic variables have been used as explanatory variables in numerous studies of the determinants of risk preferences (for example, Guiso and Paiella 2008, Yesuf and Bluffstone 2009, Tanaka et al. 2010)

Table 3 indicates that the average age of the experiment participants (45 years) is less than that of the ERHS respondents (51 years). For both the experiment and the ERHS, approximately 30% of respondents are female, and the average household has 6 members. A slightly higher fraction of round 7 ERHS respondents (74%) report farming as their primary occupation, as compared to the experiment participants (66%) – these high proportions of farmers in the samples highlight the importance of agricultural production in this region. Indeed, as noted by Hill et al. (2011), agricultural production is by far the most important source of income for ERHS households, and a majority of households in rural Ethiopia are involved in farming.

Almost all of the ERHS respondents are household heads, while only 70% of experiment participants are the heads of their respective households – this is because several attempts were made by ERHS enumerators to interview the head of each household in the survey sample.[38] On the other hand, the experiment enumerators asked households to send the member responsible for the financial decisions of the household – this member was sometimes the household head, but was often the most educated member of the household or the household head's spouse.

Formal schooling in the sample is relatively low – the average experiment participant had four years of formal schooling, while the average ERHS

[38] If it was not possible to interview the household head in the ERHS, the most knowledgeable person in the household was interviewed.

respondent received only two years of formal schooling.[39] A number of studies have found a significant relationship between risk attitudes and cognitive ability, with most obtaining evidence of a negative relationship between risk aversion and cognitive skills (for example, Burks et al. 2009, Dohmen et al. 2010, Dave et al. 2010). However, years of schooling may be a poor proxy for the ability to solve the mathematical problems that rural inhabitants encounter in everyday life (Cole et al. 2009). Therefore, in line with the work of Cole et al. (2009) and Hill et al. (2011), who analyze the determinants of risk aversion and microinsurance take-up in rural India and Ethiopia (respectively), I also include a direct measure of quantitative literacy to capture mathematical skills and cognitive ability. Quantitative literacy is measured as the fraction of 6 questions – which vary between the ERHS and the experiment – assessing probability and mathematical skills answered correctly (see Table 9 in Appendix A for more details on these questions).[40] Table 3 shows that the quantitative literacy measure is slightly below 60% for both the round 7 ERHS respondents and the experiment participants, on average.[41,42] Dave et al. (2010) note that such direct measures of quantitative literacy appropriately capture mathematical skills and the proficiency with numbers in everyday life, and these have a distinct – and possibly larger – impact on risk attitudes than general intelligence; therefore, it is important to distinguish between quantitative literacy and other measures of cognitive ability,

[39] This substantial difference in the average years of formal schooling can be attributed to the fact that, as noted above, experiment participants were often the most educated members of their households, while ERHS respondents were household heads (who were not necessarily the most educated members of their households).

[40] Hill et al. (2011) use the same measure of quantitative literacy in their analysis of the demand for weather indexed microinsurance using the ERHS dataset.

[41] The final quantitative literacy question listed in Table 9 asks whether it is riskier for farmers to plant one crop or multiple crops. While it is generally expected that diversification reduces risk exposure (and hence planting multiple crops is considered as the correct answer to this question in the formulation of the quantitative literacy variable), in certain cases farmers in poor rural economies can improve risk coping through specialization, that is, by planting a single crop which is resistant to pests, droughts and other environmental risk factors (Fafchamps 2003). For example, Fafchamps (2003) notes that in many areas of West Africa, millet is the only cultivated crop, and specialization in the cultivation of this single, robust crop is less risky and forms the main source of income smoothing. Therefore, the correct answer to this particular quantitative literacy question is not immediately clear. However, the results for all the specifications in this chapter remain substantively the same when the quantitative literacy measure excludes this question.

[42] Cole et al. (2009), on the other hand, find that respondents in their sample from rural Gujarat (India) correctly answered only 34% of similarly-framed quantitative literacy questions.

such as years of formal education.

3.1 Instrumental variables

As noted in earlier sections, basic interval regression and structural maximum likelihood estimates of the determinants of risk preferences may be biased and inconsistent, due to the endogeneity associated with wealth and background risk in these specifications. In Sections 2.2 and 2.3, I outlined estimation strategies involving instrumental variables that would help overcome these endogeneity issues and consistently estimate the impact of these variables on risk preferences.[43] While some studies do account for the endogeneity of wealth when analyzing the determinants of risk preferences (for example, Tanaka et al. 2010), only Guiso and Paiella (2008) account for the endogeneity of both wealth and background risk and jointly estimate their impact on risk preferences. Guiso and Paiella (2008), however, stress the difficulty in finding suitable IVs for wealth and background risk when estimating the determinants of risk preferences, and only use EUT specifications in which they instrument for wealth, while using a regional-level exogenous variable to proxy for background risk. I, on the other hand, consider household-level measures of both wealth and background risk, and instrument for both these variables in order to account for endogeneity and consistently estimate causal effects – to my knowledge, this is the first analysis of the determinants of risk preferences in a low-income setting which involves instrumenting for both wealth and background risk.

As IVs for household wealth, I use the following variables from the ERHS: the land inherited by the household, the distance from the household's village to the nearest town interacted with the gender of the respondent, and the total number of households in the village (village size) interacted with the gender of the respondent. Deininger and Jin (2006) and Dercon and Ayalew (2007)

[43] Tanaka et al. (2010) note that while an ideal methodology to overcome these endogeneity issues would involve the randomized assignment of individuals to economic circumstances, the use of IVs provides an appropriate, feasible alternative.

note that land rights, and their transferability, in rural Ethiopia remains highly restricted – the market for land is virtually non-existent, and transfer of land rights through sale or exchange is severely limited. Therefore, inheritance to immediate family members is the primary channel through which land rights are transferred (Deininger and Jin 2006), and hence the land inherited by the household is expected to be a crucial determinant of the amount of land the household currently has rights over. Further, with agricultural production as the primary source of income in this region, households with more inherited land – and thus more land available for farming – are expected to be richer and have greater resources to invest in livestock, which is an important marketable asset in this region (Rosenzweig and Binswanger 1993, Dercon 2004).[44] Additionally, Rosenzweig and Binswanger (1993), in their study of weather risk and agricultural investment in rural India, note that the wealth inherited by an individual is orthogonal to his preferences for risk. Therefore, in the context of risk preferences in rural Ethiopia – and the specifications outlined earlier – inherited land is an exogenous variable that is an important determinant of a household's wealth trajectory, and is expected to only impact risk preferences through its impact on both wealth measures used in this study (and in particular, the total land owned by the household). Thus, it is expected to satisfy the exclusion restriction (for IVs). Further, Foster and Rosenzweig (1995) and Munshi and Rosenzweig (2009) use inherited assets as IVs for household wealth in their analyses of technology adoption and migration, respectively, in rural India.

Drawing on the econometric strategy of Fafchamps et al. (1998), I also use gender, a demographic characteristic that is not endogenously determined by risk preferences, interacted with village-level variables – distance to the nearest town and village size – as IVs for wealth. While many studies find that risk preferences vary by gender (for example, Eckel and Grossman 2008, Guiso and Paiella 2008) – as a result of which gender is included as a determinant of risk attitudes – there is little reason to suspect that this correlation varies

[44] Fafchamps et al. (1998) note that livestock transactions are widespread in rural Africa.

by village and village-level factors such as distance to the nearest town and village size. Therefore, these interaction terms are not expected to directly affect risk attitudes and thus can be excluded as determinants of risk preferences (that is, they are expected to satisfy the exclusion restriction). Further, village characteristics such as distance to the nearest town and village size are expected to affect household wealth. For example, villages containing more households could imply limited land resources and less land ownership per household, or indeed even more land ownership per household, if villages with more households occupy disproportionately larger geographical areas; Walker and Ryan (1990), analyzing data from rural India, find evidence that household wealth varies significantly with village size. Additionally, living in a village relatively close to a town could increase household wealth by providing access to larger markets for goods and services – Tanaka and Munro (2012) note that the distance of a village to the nearest town is likely to affect the economic circumstances of households in that village.

Furthermore, the influence of these village-level characteristics on household wealth is expected to vary by the gender of the respondent, who in most cases is the household head – and thus in control of household finances – or, in the case of the Ethiopian experiment, the member who makes the financial decisions of the household. Male and female respondents could account for these village-level factors differently in household financial decision-making, implying a differential impact (by gender) of these factors on household wealth – indeed, Powell and Ansic (1997), using experimental data, find that males and females adopt different strategies in financial decision-making, and these differences are not explained by contextual factors but instead are likely to represent general traits. Thus, these interaction terms are expected to impact household wealth, and influence risk preferences only through their effect on wealth.[45] Along similar lines, Guiso and Paiella (2008) also use interactions involving the size of the town of residence as IVs for wealth in their analysis of the determinants of risk preferences using Italian survey data.

As an instrument for background risk (measured in this study by the inter-temporal standard deviation of consumption), I use the inter-temporal standard deviation of annual rainfall in the village over the 15-year period 1994-2009, interacted with the gender of the respondent.[46,47] Agricultural production is by far the most important source of income for ERHS households – households in the survey derive 80% of their income from agriculture, on average, while 94% of households are heavily dependent on the main Kiremt rains for agricultural production and income (Hill et al. 2011). Additionally, as noted by Dercon et al. (2011), rainfall variability and severe droughts are highly prevalent in rural Ethiopia, and rainfall risk is a persistent concern and a major threat for inhabitants. Thus, though rainfall in rural Ethiopia is low and erratic (with drought a frequent occurrence) and agricultural output is heavily dependent on the level of rainfall, agricultural production is the primary income-generating

[45] Though the experiment enumerators asked households to send the member responsible for the financial decisions of the household, in this East African context, unless households are female-headed, it is possible that women are not the financial decision-makers; in other words, it is possible that some of the women involved in the experiment were not actually the financial decision-makers of their households. However, if this is the case, the instruments are still expected to be valid, but the joint F-statistic of the excluded instruments in the first stage regression would provide the lower bound for the explanatory power of the instruments on the endogenous variables; further, since these F-statistics (reported in Appendix B) are quite large in magnitude, there is evidence that the instruments are relevant and have significant explanatory power for both measures of wealth (total livestock units and total land owned) and for background risk (measured by the standard deviation of consumption).

[46] The rainfall data for the years 1994-2004 is obtained from the Ethiopian National Meteorology Agency (since access to this data is only available up to 2004), and for the years 2004-2009 is obtained from the Livelihoods, Early Assessment and Protection (LEAP) project database. While the data from Ethiopian National Meteorology Agency provides information on rainfall measured at the weather stations nearest in distance to the 15 villages in the sample, the LEAP rainfall data is only available at the woreda (district) level. Therefore, to combine the LEAP data with the Ethiopian National Meteorology Agency data, I use the GPS coordinates of the 15 weather stations nearest to the sample villages to obtain rainfall information from the LEAP database for those coordinates – this is identical to the strategy adopted by Hill et al. (2011), who also use rainfall data from the LEAP database in their analysis of risk aversion and index insurance take-up in rural Ethiopia.

[47] The 15-year period 1994-2009 corresponds to the period over which the seven rounds of the ERHS were conducted, and thus the period over which the inter-temporal standard deviation of household consumption is measured.

activity for a majority of households in the region – as a result, there are sizable fluctuations in agricultural income and hence household consumption, mainly due to the large variation and uncertainty in rainfall (Dercon 2004, Porter 2011).

Dercon (2004) and Porter (2011) find that while households in this region are, to a large extent, able to smooth consumption in the face of idiosyncratic shocks, household consumption remains extremely vulnerable to rainfall fluctuations. This is because unlike idiosyncratic risk, weather risk is spatially covariant (that is, covariate in nature), and thus rural inhabitants cannot utilize informal risk-sharing arrangements – which generally occur intra-village between geographically proximate households – to insure against rainfall fluctuations that impact all households in a local environment (such as a village) (Rosenzweig and Binswanger 1993). Further, there are few formal risk-coping options available to these rural inhabitants (Fafchamps 2003). Thus, as noted by Rosenzweig and Binswanger (1993), weather risk is the factor contributing to income variability that has the largest (adverse) impact on consumption and welfare, since rural inhabitants are extremely vulnerable to weather risk, which usually remains largely uninsured; in addition, Fafchamps et al. (1998) find that deviations in rainfall explain most of the fluctuations in income caused by weather risk. Therefore, the standard deviation of rainfall identifies variation in income, and thus consumption, that is exogenous and difficult to anticipate or insure against (Rosenzweig and Binswanger 1993, Fafchamps et al. 1998, Porter 2011).

Once again, the impact of rainfall variation on household background risk is likely to vary by the gender of the respondent, as male and female respondents may differently account for rainfall variation in the financial decisions, production strategies and investment choices that determine the background risk faced by their households (Powell and Ansic 1997). Further, the effect of gender on risk preferences is not expected to vary by the standard deviation of village rainfall – that is, the interaction between rainfall variation and gender is not expected to influence risk preferences directly but only through its impact on background risk, thus satisfying the exclusion restriction for IVs.

Indeed, Fafchamps et al. (1998) also use interactions between rainfall deviation and demographic characteristics to capture exogenous variation in income.

It is important to note that the exogeneity, and thus the validity, of the three interaction IVs involving village-level variables could be called into question if risk preferences drive the widespread migration of households from one region to another. However, Dercon and Hoddinott (2004), in their introduction to the ERHS dataset, report a very low rate of migration among ERHS households. This is because agricultural production is the primary occupation for a majority of the households, and due to the rigid and incomplete system of land rights in rural Ethiopia, the transfer of land rights through sale or exchange is severely limited – this implies that households do not have rights to land (or land use) in other regions, and cannot purchase land (or land rights) to farm in those regions (Dercon and Hoddinott 2004, Dercon and Ayalew 2007). Deininger and Jin (2006) note that this prevents migration, as farmers, in most cases, have rights only to inherited land, and cannot obtain rights to land (or land use) in other villages. Indeed, Ezra and Kiros (2001) find that marriage is the primary reason for voluntary migration, with women migrating to live with their husbands – migration for economic reasons, and thus due to risk preferences, is negligible. Additionally, since 1975, much of the internal migration in Ethiopia has been due to government policies involving forced relocation, resettlement and land redistribution, and is thus exogenous to risk preferences (Deininger and Jin 2006, Fransen and Kuschminder 2009). Therefore, since a majority of the migration in rural Ethiopia over the last three decades has been forced, and voluntary migration is in most cases not driven by economic or risk factors, it is reasonable to assume that the interaction IVs are valid, since most of the migration in rural Ethiopia is likely to be exogenous to risk preferences.[48,49]

[48] For more on this issue, see Section 5.

[49] Even though household migration is not common in rural Ethiopia, there remains the possibility of temporary migration (that is, some members of the household absent for a limited period of time).

4 Results of Empirical Analysis

As mentioned earlier, the results in Chapter 1 indicate that EUT-CRRA is likely to provide an appropriate description of the decision-making process generating the choices in the Binswanger-style lottery of the Ethiopian experiment; on the other hand, the choices in the insurance decision problems are better described by a non-EUT decision model, the two-part power utility specification with loss aversion (but no probability weighting).

Keeping these findings in mind, the empirical analysis is as follows. I first analyze the choices in the Binswanger-style lotteries of the Ethiopian experiment and ERHS using the asset integration model described in Section 2.1, which assumes EUT-CRRA preferences, in order to evaluate whether participants' choices in these decision problems reflect narrow framing or asset integration. Next, I utilize the basic interval regression, two-step interval regression and LIML estimators described in Section 2.2 (in conjunction with the instrumental variables described in Section 3.1) to analyze choices in the Binswanger-style lotteries of the Ethiopian experiment and ERHS and evaluate the determinants of risk aversion, assuming EUT-CRRA utility. Finally, I use the structural maximum likelihood estimator and two-step structural maximum likelihood estimator described in Section 2.3 – which are appropriate for non-EUT decision models – to analyze the choices in the four insurance decision problems of the Ethiopian experiment, assuming that the decision-making process generating choices in these problems is given by the two-part power utility function with no probability weighting and a reference point of approximately 17 birr.[50,51] The results of the empirical analysis are presented below.

4.1 Asset integration

Table 4 presents the results for the maximum likelihood estimation of the asset integration model detailed in Section 2.1, using data from the Binswanger-style lottery of the Ethiopian experiment (benchmark decision problem B), and total livestock units as the measure of outside wealth.

When I assume the argument of the EUT-CRRA utility function to be $\delta W + x$ (Panels B and C), the estimated coefficient of relative risk aversion is around 0.5 – this implies that participants in the experiment exhibit moderate risk aversion in B, on average, and the estimates are similar to the estimate of average r obtained when only x is assumed to be the argument of the utility function (Panel A). Additionally, these estimates are similar to the coefficient of relative risk aversion of around 0.5 recently estimated for Ethiopian, Ugandan and Indian subjects in a traditional laboratory experiment (framed in the abstract) conducted by Harrison et al. (2010); Botelho et al. (2005) also estimate average r to be approximately 0.6 when analyzing experimental data from Timor-Leste.

Panel B of Table 4 shows that when integrated wealth δW is considered

[50] I was unable to conduct the asset integration analysis for choices in the insurance decision problems of the Ethiopian experiment, as these choices are best characterized by non-EUT decision models involving more than one preference parameter, and the optimization routine did not converge when I attempted to estimate integrated wealth in addition to the preference parameters of these models. Additionally, as noted in Chapter 1, this may be a result of the lower precision and power provided by the Ordered Lottery Selection elicitation procedure.

[51] It is important to note that, in the Ethiopian experiment, the benchmark decision problem was framed very differently from the insurance decision problems. These important differences in the framing and nature of the benchmark and insurance problems are likely to lead to preference reversals on the part of subjects, which cause a systematic change in decisions (and the decision-making process) in otherwise identical tasks – thus, it is possible (even likely) that an individual makes decisions in these problems according to different decision models. As noted in the previous paragraph, in Chapter 1, I do indeed find evidence of these preference reversals – the evidence indicates that different decision models may be generating choices in the benchmark and insurance problems (even for the same individual), and thus choices in these problems should be considered separately in estimations. Therefore, in this chapter, I continue to analyze the choices in the benchmark and insurance problems separately (as done in Chapter 1), assuming different decision-making processes.

Table 4. *Maximum likelihood estimates of asset integration model: Experiment Binswanger-style lottery*

Coefficient	Estimate	Standard Error	p-value	Lower 95% Confidence Interval	Upper 95% Confidence Interval	Observations
A. Utility function: $U(x) = \dfrac{x^{1-r}}{1-r}$						
r	0.537	0.025	0.000	0.488	0.586	378
B. Utility function: $U(W,x) = \dfrac{(\delta W + x)^{1-r}}{1-r}$; δW = Constant						
r	0.470	0.175	0.007	0.127	0.813	378
δW (constant)	66.59	230.6	0.773	-385.4	518.6	378
C. Utility function: $U(W,x) = \dfrac{(\delta W + x)^{1-r}}{1-r}$; W = Total livestock units						
r	0.530	0.031	0.000	0.471	0.590	372
δ	0.699	1.565	0.655	-2.368	3.768	372

The log-likelihoods corresponding to the estimations in Panels A, B and C are -629.4, -628.9 and -619.1. Wald test p-values and 95% confidence intervals reported.

as a constant argument of the utility function, its estimate is statistically indistinguishable from zero – thus, using a Wald test, the hypothesis of narrow framing ($\delta W = 0$) cannot be rejected at the 10% significance level.[52] Further, the δW estimate of approximately 65 birr (5 US dollars) is extremely small in magnitude, compared to the total value of wealth of a respondent's household. Even though the standard error is quite large, the upper limit of the 95% confidence interval (518 birr) itself represents only a minute fraction of the total value of household wealth – this is highlighted by the fact that average monthly household consumption for ERHS households, measured in round 7 of the survey, is 1100 birr (as mentioned in Section 3).[53] Therefore, this provides evidence that participants in the experiment consider the argument of the utility function under EUT-CRRA to be only the lottery prize in the decision problem (x), and make decisions by taking into account only the utility from lottery prizes; in other words, participants' choices in the experiment Binswanger-style lottery reflect narrow framing, rather than the integration of outside assets. This is in line with the results of Harrison et al. (2007) and Heinemann (2008), who estimate a similar asset integration model (assuming EUT-CRRA utility) and obtain estimates of the constant δW term that are relatively small in magnitude, thus finding no evidence of asset integration among samples of experiment participants in the United States.

Panel C of Table 4 presents the results when the δ parameter is estimated separately (in addition to r), and total livestock units are used as the measure of wealth.[54] The estimate of δ is statistically indistinguishable from 0, and this

[52] In this study, I also conducted likelihood ratio tests in addition to Wald tests to test the various hypotheses – in all cases the conclusions from both tests are identical.

[53] The relatively wide confidence interval of the estimate of integrated wealth δW may reflect the diverse demographic backgrounds of the respondents, as noted by Andersen et al. (2006b), or it could be a manifestation of the relatively small number of observations used in the estimation.

[54] When I attempted to maximize the log-likelihood function with respect to parameters δ and r, using total land owned as the measure of wealth, the optimization routine did not converge.

once again is consistent with narrow framing, as opposed to asset integration.[55,56]

Thus, the estimates in Panels B and C provide some indication that the decisions of participants in the Ethiopian experiment Binswanger-style lottery reflect narrow framing, rather than asset integration. This inference is strengthened by the fact that the estimate of r obtained is similar for the models with and without the inclusion of the integrated wealth term.

Table 5 presents the estimation results for the asset integration model using the choices in the Binswanger-style lottery in the ERHS. Once again, I find that the estimates of r imply risk aversion among respondents, and are similar in magnitude for the models with and without the inclusion of integrated wealth. However, the estimated coefficient of relative risk aversion (r) of around 0.25 is lower than that obtained for the decisions in the Ethiopian experiment Binswanger-style lottery (approximately 0.5) – this indicates that individuals exhibit lower risk aversion in the ERHS lottery as compared to the experiment lottery.[57] This difference in estimated risk aversion could be because the two samples have different preferences over risk. Even though the experiment households were chosen from within the larger sample of ERHS households, the ERHS sample (1,350 households) is substantially larger than – and thus considerably different from – the experiment sample (378 households); additionally, it is not necessarily the case that, for a particular household, the respondent in the ERHS survey is also the experiment participant. Alternatively, the difference in estimated risk aversion could be a result of the difference in context of the experiment and ERHS lotteries, which may impact elicited risk

[55] In accordance with the strategy used by Andersen et al. (2006a), the maximum likelihood estimation actually provides estimates of a transform of δ, to ensure that $\delta \in (0, 1)$ as required by the model described in Section 2.1. Specifically, I obtain estimates of δ', where $\delta = 1/(1 + \exp(\delta'))$. This non-linear function of δ' can be easily calculated from the estimates, and the "delta method" can be used to provide estimates of the standard error and p-value (Oehlert 1992). Thus, $\delta = \frac{1}{2}$ when $\delta' = 0$, $\delta > \frac{1}{2}$ when $\delta' < 0$ and $\delta < \frac{1}{2}$ when $\delta' > 0$; $\delta' \in (-\infty, +\infty)$, so $\delta \in (0, 1)$.

[56] However, the standard error of the δ estimate is quite large, possibly due to the relatively small number of choices used in this estimation (372); as a result, the hypothesis of full asset integration ($\delta = 1$) also cannot be rejected at the 10% significance level using a Wald test (or likelihood ratio test).

[57] Using a Wald test, the hypothesis that r is equal in both cases is rejected at the 1% significance level.

Table 5. *Maximum likelihood estimates of asset integration model: ERHS Binswanger-style lottery*

Coefficient	Estimate	Standard Error	p-value	Lower 95% Confidence Interval	Upper 95% Confidence Interval	Observations
Utility function: $U(x) = \frac{x^{1-r}}{1-r}$						
r	0.242	0.015	0.000	0.212	0.271	1348
Utility function: $U(W,x) = \frac{(\delta W + x)^{1-r}}{1-r}$; δW = Constant						
r	0.244	0.016	0.000	0.213	0.276	1348
δW (constant)	0.029	0.032	0.356	-0.033	0.091	1348
Utility function: $U(W,x) = \frac{(\delta W + x)^{1-r}}{1-r}$; W = Total livestock units						
r	0.282	0.037	0.000	0.210	0.355	1348
δ	0.039	0.055	0.478	-0.069	0.147	1348

The log-likelihoods corresponding to the estimations in Panels A, B and C are -2076.8, -2076.3 and -2169.5. Wald test p-values and 95% confidence intervals reported.

preferences (as found in Chapter 1) or the degree of asset integration; though the gain frame Binswanger-style lottery is similar in both cases (and framed in the abstract), the ERHS subjects were interviewed in their own homes (as part of the survey), while the experiment was conducted in a lab setting at external sites.

Panel B of Table 5 shows that the estimate of integrated wealth δW is statistically indistinguishable from 0, implying that the hypothesis of narrow framing ($\delta W = 0$) cannot be rejected at the 10% significance level (using a Wald test). Further, the estimated value of integrated wealth (0.029) is extremely small in magnitude, once again indicating that there is no asset integration, and participants behave as if maximizing EUT-CRRA utility over only the lottery prizes; this is in line with the results of Harrison et al. (2007) and Heinemann (2008). These results are echoed in Panel C, which provides estimates of the parameters δ and r, with total livestock units used as the measure of wealth.[58] The estimate of δ (0.039) is both small in magnitude and statistically indistinguishable from 0, again providing evidence of narrow framing, as opposed to asset integration.[59] This estimate implies that a weight of only 3.9% is placed on outside wealth in the argument of the utility function, indicating that participants only consider the utility from prizes in each lottery (Andersen et al. 2006a). Andersen et al. (2006a), analyzing the choices of university members in the United States – within an EUT-CRRA framework – also obtain a similar estimate of the δ parameter (0.03), which is statistically indistinguishable from 0, implying that virtually no weight is placed on outside wealth in the argument of the utility function. Additionally, in similar estimations using experimental data from Denmark, Harrison et al. (2011) obtain estimates of δ that are very close to 0.

Thus, the results from both the ERHS and Ethiopian experiment Binswanger-style lotteries are consistent with the hypothesis of narrow framing – there is evidence that subjects consider only the lottery prize x as the argument of the EUT-CRRA utility function, and do not integrate outside wealth when making decisions in these lotteries. That is, participants in both the ERHS and experiment seem to isolate decisions in the lottery from other decisions in their lives, and make choices by considering only the prospective gains and losses associated with the current decision, independent of outside wealth (Heinemann 2008). This is in line with the results of numerous experimental studies involving subjects from both developed economies (for example, Andersen et al. 2006a, Harrison et al. 2007, Heinemann 2008) and from low-income countries (for example, Binswanger 1981, Schechter 2007, Fafchamps et al. 2013). Indeed, Heinemann (2008) notes that the existing evidence for narrow framing, as opposed to asset

[58] Once again, when I attempted to maximize the log-likelihood function with respect to parameters δ and r, using total land owned as the measure of wealth, the optimization routine did not converge.

[59] Unlike for the data from the Ethiopian experiment Binswanger-style lottery, the standard errors of the δ and δW estimates are relatively small in this case; additionally, the hypothesis of full asset integration ($\delta = 1$) is rejected at the 1% significance level (using a Wald test). The small standard errors and corresponding narrow confidence intervals are likely due to the substantially greater number of choices in the ERHS lottery compared to the experiment lottery.

integration, is strong.[60] He suggests that one possible reason for the prevalence of narrow framing in experimental decision-making is mental accounting, whereby subjects treat each experimental decision problem as a single entity for which they have an aspirational level that they try to achieve.

In accordance with the results in this section, I consider the argument of utility to be only the lottery prize x for all future estimations in this chapter (which involve evaluating the determinants of risk preferences).

4.2 Determinants of risk aversion in an EUT framework

Table 6 presents the results for the different interval regression estimations of the determinants of risk aversion using the choices from the Binswanger-style lottery of the Ethiopian experiment, and assuming EUT-CRRA utility. Columns (1) and (2) provide the results for the basic interval regression model, while Columns (3)-(6) provide the results for the two IV interval regression models – the results for the two-step specification are displayed in Columns (3) and (4), while those for the LIML model are displayed in Columns (5) and (6).[61]

Column (1) of Table 6 shows that total livestock units are negatively related to the coefficient of relative risk aversion r – using the basic interval regression estimator, the coefficient of total livestock units is estimated to be negative and statistically significant at the 10% level. Further, the coefficient estimate implies that an increase in total livestock units by one decreases r by 0.11.[62] On the other

[60] However, Heinemann (2008) also notes that while the evidence for non-integration of outside wealth is strong, some studies do find evidence of integration of laboratory income from prior decision problems within the experiment (labelled "local asset integration"), in experiments where participants are allowed to accumulate wealth over consecutive decision problems (for example, Andersen et al. 2006a, Andersen et al. 2008, Fafchamps et al. 2013). However, in the Ethiopian experiment, each decision problem (and the corresponding payment to participants) was independent, and participants played for real money – and thus were paid for – only one of the three decision problems they responded to. Further, the decision problem to be played for real money was randomly selected only at the end of the session, and subjects did not know their monetary payoff after each problem. Thus, with such an experimental setup, wealth accumulation from previous decision problems is highly unlikely (Charness et al. 2013).

[61] Following Bettin and Lucchetti (2010), robust standard errors are reported for the all the estimations in this chapter.

[62] As mentioned in Section 3, one total livestock unit is equivalent to one oxen or ten goats (or sheep).

Table 6. *Determinants of risk aversion: Experiment Binswanger-style lottery*

	Basic Interval Regression		Two-step		LIML	
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Total livestock units	-0.111*		0.293		0.306	
	(0.0593)		(0.197)		(0.195)	
Total land owned		0.0104		0.357		-19.77
		(0.00994)		(0.295)		(25.19)
Std. dev. of consumption	0.00126**	0.00111**	0.000916	-0.000217	-0.00104	0.0226
	(0.000576)	(0.000566)	(0.00174)	(0.00198)	(0.00182)	(0.125)
Household size	-0.228**	-0.263***	-0.269**	-0.321	-0.267*	5.692**
	(0.0895)	(0.0909)	(0.134)	(0.216)	(0.137)	(2.776)
Age	0.0203	0.0189	0.0195	0.0173	0.0195	0.223
	(0.0137)	(0.0136)	(0.0147)	(0.0158)	(0.0147)	(0.154)
Gender	-0.890	-0.0834	-0.625	-0.727	-0.615	-6.923
	(0.547)	(0.546)	(0.609)	(0.588)	(0.613)	(6.558)
Formal schooling	0.0266	0.0306	0.0495	0.0396	0.0505	0.0926
	(0.0462)	(0.0462)	(0.0508)	(0.0498)	(0.0510)	(1.120)
Quantitative literacy	-1.635*	-1.542*	-1.110	-1.324	-1.087	-6.041
	(0.927)	(0.921)	(0.260)	(0.973)	(1.049)	(12.830)
Household head	-0.227	-0.223	-0.233	0.175	-0.231	-20.28*
	(0.425)	(0.429)	(0.456)	(0.752)	(0.458)	(12.08)
Farmer	-0.812	-0.859*	-0.887	-1.184*	-0.892	16.58
	(0.517)	(0.520)	(0.639)	(0.638)	(0.647)	(18.04)
Constant	5.001***	4.025***	1.466	3.699**	-1.351	24.22
	(1.575)	(1.483)	(2.210)	(1.587)	(2.267)	(20.68)
Observations	369	369	365	365	365	365
Log-likelihood	-806.6	-807.9	-792.2	-795.7	-791.9	-794.0
Wald test for exogeneity			6.376	1.980	5.684	5.911
Wald test p-value			0.0413	0.372	0.0583	0.0520

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Dependent variable is the range of the coefficient of relative risk aversion. Standard errors are in parentheses.

hand, Column (2) shows that the relationship between total land owned and risk aversion is not statistically significant.

Columns (1) and (2) also indicate that background risk is positively related to r, in line with Gollier and Pratt's (1996) risk vulnerability hypothesis – the coefficient of the inter-temporal standard deviation of consumption is estimated to be positive and statistically significant at the 5% level, both when total livestock units and total land owned are used as the measure of wealth. Additionally, the results also indicate that household size and quantitative literacy are negatively related to risk aversion.

However, as mentioned in earlier sections, the coefficient estimates from the basic interval regression may be biased and inconsistent due to the endogeneity associated with wealth and background risk. To overcome these endogeneity issues and obtain consistent estimates, I utilize instrumental variables (listed in Section 3.1) in combination with the two-step and LIML estimators detailed in Section 2.2.

The first stage of the two-step procedure is identical to that of the IV Two-Stage Least Squares procedure, and involves OLS regressions of the endogenous variables – wealth and background risk – on the IVs and exogenous determinants of risk preferences (see Section 2.2). Table 10 in Appendix B presents the results of these first stage regressions for data from the experiment Binswanger-style lottery. The first stage results in this table indicate that the instruments are relevant and have significant explanatory power for both measures of wealth (total livestock units and total land owned) and for background risk (measured by the standard deviation of consumption) – the joint F-statistic of the excluded instruments is approximately 33, 16 and 13 for total livestock units, total land owned and the standard deviation of consumption, respectively. This implies that the p-value for the F-test of joint significance of the excluded instruments is very close to zero in all three cases.

Columns (3) and (4) in Table 6 present the results for the second stage interval regression of this two-step procedure, with the range of r implied by the participant's choice in the experiment Binswanger-style lottery as the dependent variable. The results indicate that neither of the two wealth measures has a statistically significant impact on risk aversion. These results are echoed in Columns (5) and (6), which show that the estimated impact of wealth on risk aversion is statistically insignificantly different from 0, when using the LIML estimator. Thus, unlike the basic interval regression estimates, the results from the two-step and LIML estimations are consistent with the inference that wealth does not affect risk aversion. This is in line with the experimental results of Binswanger (1980) (India), Mosley and Verschoor (2005) (Ethiopia, Uganda and India) and Falco (2012) (Ghana), who also find no significant relationship between wealth and risk preferences. Indeed, Zyphur et al. (2009) find that risk preferences are largely heritable, and thus may not be affected by current wealth or background risk. Yesuf and Bluffstone (2009), on the other hand, analyzing the experimental choices of subjects from farm households in Northern Ethiopia, find that both livestock and land ownership are negatively related to risk aversion; however, while they attempt to address potential endogeneity due to time-invariant factors, their results may be biased and inconsistent as they do not instrument for wealth in their specifications and do not appropriately account for possible endogeneity due to time-variant factors.

The results from the two-step and LIML estimations – Columns (3) to (6) – also indicate that background risk no longer has a statistically significant impact on risk aversion exhibited in the experiment Binswanger-style lottery. Additionally, the coefficient estimates of background risk are extremely small in magnitude, lending further credence to the hypothesis that background risk does not affect the risk aversion of participants in the experiment lottery. Lusk and Coble (2008), analyzing the influence of experimentally-induced background risk among university members in the United States, reject the risk vulnerability hypothesis of Gollier and Pratt (1996) and find evidence of a negligible impact of background risk on risk preferences. Additionally, Alessie et al. (2002), analyzing the portfolio structure of Dutch households using survey data, find no significant relationship between the demand for risky assets and background income uncertainty. However, these results are in contrast to those of Harrison et al. (2007) and Guiso and Paiella (2008), who find that experimentally-induced and non-experimental background risk, respectively, significantly increase risk aversion.[63]

Very few of the individual-level and household-level demographic characteristics have a statistically significant impact on risk aversion in the two-step and LIML estimations; Guiso and Paiella (2008) also find this to be the case after controlling for the effect of wealth and background risk. The results regarding household size are mixed – while the IV estimations involving livestock ownership (Columns (3) and (5)) indicate a negative and statistically significant impact of household size on risk aversion, the LIML estimation involving land ownership (Column (6)) yields a positive and statistically significant (at the 10% level) coefficient estimate of household size. Column (4) also shows that farmers are less risk averse than non-farmers, while Column (6) shows that household heads have lower aversion to risk than other household members, *ceteris paribus*. The latter finding may be because those subjects who were answerable to their household heads – and probably had to give most (or all) of their experimental earnings to their household heads – were averse to returning with little (or no) money from the experiment (other than the participation fee) and thus were more likely to choose safer options in the decision problems.

Table 6 also reports the results of Wald tests for (joint) exogeneity of wealth and background risk, when the two-step and LIML estimators are utilized.

[63] The results of Guiso and Paiella (2008), however, are based on a single hypothetical survey question – problems relating to the use of questions involving hypothetical payoffs to elicit risk preferences have been well documented (see, for example, Holt and Laury 2002, Harrison and Rutström 2008a). Indeed, Harrison and Rutström (2008b) note that preferences elicited using hypothetical questions are significantly different from those elicited using questions involving real economic payoffs, and thus decisions in hypothetical questions may not provide meaningful information on economic decisions in real life; Charness et al. (2013) also cite evidence from numerous studies indicating that risk preferences elicited in non-incentivized tasks do not reflect true risk attitudes, noting the importance of incentive-compatibility in eliciting risk preferences.

This test involves testing the null hypothesis that wealth and background risk are both exogenous – that is, $\mathbf{\Omega} = \mathbf{0}$ in Equation (18) – using a Wald test (Bettin and Lucchetti 2010).[64] These Wald tests indicate that for the specifications in Columns (5) and (6), the null hypothesis of exogeneity of wealth and background risk is rejected at the 10% level, while this hypothesis is rejected at the 5% level for the specification in Column (3). This provides an indication that wealth and background risk are indeed endogenous in the context of risk preferences; thus, the basic interval regression estimates are likely to be biased and inconsistent. Additionally, the coefficient estimates differ substantially between the basic and IV interval regression specifications, lending further credence to this hypothesis and highlighting the importance of utilizing instrumental variables for drawing accurate causal inferences regarding the determinants of risk preferences. Further, likelihood ratio tests for over-identifying restrictions of the instruments yield test statistics with p-values greater than 0.10 for the two-step and LIML specifications – therefore, the null hypothesis of exogeneity of the instruments cannot be rejected at the 10% level for all four IV specifications. This provides some evidence for the validity of the instruments used to address endogeneity (Bettin and Lucchetti 2010).

I now turn my attention to the results, provided in Table 7, of the estimations involving choices in the ERHS Binswanger-style lottery. It is important to note that most studies which analyze the determinants of risk preferences using data from decision problems in experiments (with relatively small subject pools) do not include village or other location dummies as explanatory variables (for example, Wik et al. 2004, Lusk and Coble 2008, Tanaka et al. 2010). However, studies which explore the determinants of risk preferences using data from surveys involving a large number of respondents often control for the location of the respondent – for example, Guiso and Paiella (2008), who analyze risk preferences using a survey dataset of over 8000 households from across Italy, include dummies for the region of birth as explanatory variables for risk

[64] This test is in line with the exogeneity test proposed by Davidson and MacKinnon (1993).

aversion. Indeed, Tanaka and Munro (2012), when analyzing data from a large and geographically-representative sample of households in rural Uganda, find that the agro-climatic conditions of inhabitants significantly impact risk preferences. Thus, since the ERHS covers a large number of geographically-diverse households from 15 villages in rural Ethiopia – which span the primary agro-climatic zones of the country – I include village dummies as explanatory variables for risk aversion when analyzing the choices of all ERHS respondents in the ERHS Binswanger-style lottery.

Table 11 in Appendix B presents the results of the first stage OLS regressions of the two-step procedure, using data from the ERHS Binswanger-style lottery. Once again, the first stage results indicate that the instruments are relevant and have significant explanatory power for both measures of wealth and for background risk – the joint F-statistic of the excluded instruments is approximately 15, 12 and 8 for total livestock units, total land owned and the standard deviation of consumption, respectively. This implies that the p-value for the F-test of joint significance of the excluded instruments is very close to zero in all three cases. However, the explanatory power of the excluded instruments is lower than in the corresponding first stage specifications using data from the experiment Binswanger-style lottery, probably due to the addition of village dummies when analyzing the ERHS data.[65]

The likelihood ratio tests for over-identifying restrictions of the instruments yield test statistics with p-values greater than 0.10 for all two-step and LIML specifications, once again providing some indication that the instruments are valid (exogenous). Wald tests for the exogeneity of wealth and background risk also yield test statistics with p-values greater than 0.10 for all four IV interval regression specifications (see Columns (3)-(6)) – this implies that for the ERHS Binswanger-style lottery data, I am unable to reject the null hypothesis (at the 10% significance level) that wealth and background risk are both exogenous

[65] These first stage F-statistics (of excluded instruments) are still greater than those reported for most of the IV specifications of Tanaka et al. (2010), who instrument for wealth in their estimations of the determinants of risk preferences, and obtain F-statistics as low as 1.17 in some of their specifications.

Table 7. *Determinants of risk aversion: ERHS Binswanger-style lottery*

	Basic Interval Regression		Two-step		LIML	
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Total livestock units	-0.00556		0.180		0.226	
	(0.0201)		(0.255)		(0.323)	
Total land owned		-0.0147		-0.324*		-0.347*
		(0.01447)		(0.179)		(0.200)
Std. dev. of consumption	-0.000436**	-0.000428**	0.0000528	0.00156	0.000115	0.00169
	(0.000176)	(0.000175)	(0.00138)	(0.00153)	(0.00170)	(0.00162)
Household size	0.0369*	0.0387*	-0.0128	0.0362	-0.0227	0.0365
	(0.0206)	(0.0203)	(0.0594)	(0.0460)	(0.0698)	(0.0543)
Age	0.00687**	0.00707**	0.00606	0.00955**	0.00592	0.00974**
	(0.00336)	(0.00337)	(0.00389)	(0.00443)	(0.00408)	(0.00455)
Gender	-0.0115	0.000667	0.0405	0.245	0.0533	0.262
	(0.119)	(0.119)	(0.137)	(0.245)	(0.144)	(0.252)
Formal schooling	0.00299	0.00333	-0.00215	-0.0101	-0.00192	-0.0107
	(0.0185)	(0.0185)	(0.0242)	(0.0240)	(0.0267)	(0.0244)
Quantitative literacy	-0.373*	-0.363*	-0.490*	-0.189	-0.518*	-0.176
	(0.212)	(0.213)	(0.268)	(0.282)	(0.290)	(0.287)
Household head	-0.0195	-0.0227	-0.0223	-0.167	-0.0184	-0.177
	(0.170)	(0.170)	(0.177)	(0.200)	(0.180)	(0.205)
Farmer	0.190	0.198	0.123	0.323*	0.109	0.333*
	(0.122)	(0.123)	(0.143)	(0.178)	(0.152)	(0.183)
Constant	0.608	0.542	-0.795	-0.308	-1.148	-0.369
	(0.513)	(0.477)	(1.988)	(0.837)	(2.496)	(0.865)
Village dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1264	1264	1255	1255	1255	1255
Log-likelihood	-2153.6	-2153.1	-2136.1	-2134.2	-2135.9	-2134.0
Wald test for exogeneity			0.846	3.407	0.841	3.115
Wald test p-value			0.655	0.182	0.657	0.440

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Dependent variable is the range of the coefficient of relative risk aversion. Standard errors are in parentheses.

to risk preferences. However, since the endogeneity associated with these two variables is still an important concern in the context of risk preferences, I continue to focus on the two-step and LIML estimations. Additionally, the coefficient estimates differ substantially between the IV and non-IV specifications, further highlighting the issue of endogeneity.

The basic (non-IV) interval regression results in Columns (1) and (2) of Table 7 indicate that while there is no statistically significant relationship between wealth and risk aversion, background risk is *negatively* related to risk aversion – the coefficient estimate of the inter-temporal standard deviation of consumption is negative and statistically significant (at the 5% level) in both specifications.[66] However, consistent with the results from the experiment Binswanger-style lottery (Table 6), the effect of background risk is no longer statistically signifiant – and the coefficient estimate is extremely small in magnitude – in the two-step and LIML estimations. Thus, once again, there is no strong evidence that background risk impacts individuals' preferences over risk in the Binswanger-style lottery.

While the two-step and LIML estimations in Columns (3) and (5) show no statistically significant effect of livestock ownership on risk aversion, the coefficient estimate of total land owned is negative and statistically significant (at the 10% level) in Columns (4) and (6). Further, the estimated effect of total land owned on risk aversion is significant in magnitude as well. The coefficient estimates in the two-step and LIML specifications imply that an increase in total land owned by one hectare reduces r by around 0.34 – given that the average r of ERHS respondents is estimated to be only 0.25 (see Table 5), this estimated impact is significant in magnitude. This is consistent with the results of Wik et al. (2004) and Yesuf and Bluffstone (2009), who find that land ownership is negatively related to risk aversion (within an EUT-CRRA framework), using data

[66] The number of observations used in the estimations (1264 in the basic interval regression and 1255 in the two-step and LIML estimations) is less than the total number of households surveyed in round 7 of the ERHS (1348), due to missing data for some variables. However, the number of observations is very similar to that reported by Hill et al. (2011), who analyze the demand for weather indexed insurance among ERHS households and use many of the same explanatory variables.

from Zambia and Ethiopia, respectively.

Once again, the coefficient estimates of most of the demographic variables are statistically insignificant. However, the results in Columns (4) and (6) do indicate that age has a statistically significant – though small in magnitude – positive impact on risk aversion in the ERHS Binswanger-style lottery.[67] Lusk and Coble (2008) and Yesuf and Bluffstone (2009) also estimate a positive relationship between age and risk aversion; this result is also consistent with the observation of Moscardi and de Janvry (1977), who note that older individuals are generally more risk averse. Columns (4) and (6) also show that the coefficient of relative risk aversion r of farmers is around 0.3 greater than that of non-farmers, indicating that farmers are more risk averse than non-farmers (*ceteris paribus*).

While formal schooling does not have a statistically significant effect on risk aversion in both Tables 6 and 7, there is some evidence in both tables indicating that lower quantitative literacy is associated with greater risk aversion. This is in line with the results of Burks et al. (2009) and Dohmen et al. (2010), who analyze data from the United States and Germany, respectively, and find that lower cognitive ability is associated with greater risk aversion. Further, Dave et al. (2010) also find some evidence that individuals with lower mathematical skills exhibit greater risk aversion in the decision problems of their experiment (conducted in Canada). Additionally, the results could indicate that mathematical skills (captured by the quantitative literacy variable) have a distinct, and possibly larger, impact on risk preferences than formal schooling (as noted by Dave et al. 2010).

4.3 Determinants of risk preferences in a non-EUT framework

I now turn my attention to the (756) choices in the four insurance decision problems of the Ethiopian experiment, which are analyzed using the two-part

[67] The results of the basic interval regression – presented in Columns (1) and (2) – also suggest a positive relationship between age and risk aversion.

power utility specification (with a reference point of approximately 17 birr) and the structural maximum likelihood estimation technique (detailed in Section 2.3). The results of the basic (non-IV) structural maximum likelihood estimation (which may yield inconsistent estimates due to the possible endogeneity of wealth and background risk) and the two-step structural maximum likelihood estimation (incorporating instrumental variables) are presented in Table 8; the upper half of the table reports the coefficient estimates capturing the impact of the various explanatory factors on the parameter describing the curvature of the utility function (α), while the lower half reports the corresponding coefficient estimates for the loss aversion parameter (λ).[68] Tanaka et al. (2010) also only analyze the determinants of λ and α in their prospect theory specifications (and assume an exogenously-given reference point of 0).[69]

Looking first at α, the results in Table 8 indicate that wealth has a negative impact on the curvature of the utility function – the coefficient estimate of livestock ownership is negative and statistically significant at the 1% level in Columns (1) and (3), while the coefficient estimate of land ownership is negative and statistically significant at the 5% level in Columns (2) and (4). The results, however, indicate that the estimated magnitude of the impact of wealth is relatively small – the two-step estimations indicate that an increase in total livestock units by one lowers α by only 0.08, while an increase in total land owned by one hectare reduces α by only 0.1. Following Tanaka et al. (2010) and Andersen et al. (2006a), who also analyze the determinants of risk preferences using a similar two-part power utility function, I interpret lower values of α as

[68] Note that the first stage OLS specifications of the two-step structural maximum likelihood procedure are identical to those of the two-step interval regression procedure using data from the experiment Binswanger-style lottery, because each of the 378 participants in the experiment was involved in two insurance decision problems. These first stage results are presented in Table 10 in Appendix B, and indicate that the instrumental variables have significant explanatory power for the endogenous wealth and background risk variables.

[69] When I attempted to estimate the two-part power utility function without assuming the reference point (χ) to be fixed at 17 birr, but instead allowing it to be estimated jointly with the other parameters, the optimization routine did not converge. Dave et al. (2010) also encounter non-convergence when attempting to estimate the determinants of preference parameters in non-EUT decision models.

Table 8. *Determinants of risk preferences: Experiment insurance decision problems*

	Basic Structural Maximum Likelihood		Two-step	
Variables	(1)	(2)	(3)	(4)
Parameter: α				
Total livestock units	-0.0544***		-0.0756***	
	(0.0181)		(0.0242)	
Total land owned		-0.0505**		-0.105**
		(0.0201)		(0.0468)
Std. dev. of consumption	0.00000343	0.00000837	0.000112	0.000200
	(0.0000942)	(0.000106)	(0.000199)	(0.000223)
Household size	0.0206*	0.0205	0.0332**	0.0342
	(0.0105)	(0.0128)	(0.0144)	(0.0214)
Age	-0.00341	-0.00305	-0.00373	-0.00391
	(0.00275)	(0.00323)	(0.00278)	(0.00341)
Gender	-0.0350	-0.0695	-0.0109	-0.0138
	(0.0933)	(0.103)	(0.0920)	(0.101)
Formal schooling	-0.00393	-0.00675	-0.00215	-0.00247
	(0.00661)	(0.00737)	(0.00646)	(0.00705)
Quantitative literacy	-0.522***	-0.533***	-0.556***	-0.486***
	(0.140)	(0.153)	(0.145)	(0.158)
Household head	0.00796	-0.00996	-0.00332	-0.0818
	(0.0654)	(0.0691)	(0.0634)	(0.0960)
Farmer	0.0104	0.00915	0.0298	0.0691
	(0.0862)	(0.0882)	(0.0939)	(0.116)
Constant	1.366***	0.998***	1.448***	0.892***
	(0.234)	(0.199)	(0.257)	(0.196)
Parameter: λ				
Total livestock units	-0.152**		-0.755*	
	(0.0618)		(0.401)	
Total land owned		-0.0314**		-1.881***
		(0.0138)		(0.645)
Std. dev. of consumption	0.0000528	0.000301	0.00502	0.00527
	(0.00124)	(0.00125)	(0.00466)	(0.00480)
Household size	-0.161	-0.0920	-0.0569	0.287
	(0.155)	(0.188)	(0.294)	(0.347)
Age	-0.0474	-0.0459	-0.0581*	-0.0343
	(0.0314)	(0.0323)	(0.0316)	(0.0327)
Gender	-0.818	-1.289	-1.541	-1.140
	(1.251)	(1.320)	(1.326)	(1.375)
Formal schooling	-0.142	-0.178*	-0.199*	-0.142
	(0.106)	(0.105)	(0.110)	(0.107)
Quantitative literacy	-0.816	-1.149	-2.230	-1.120
	(2.266)	(2.208)	(2.450)	(2.438)
Household head	0.153	-0.159	-0.114	-2.174
	(1.160)	(1.171)	(1.235)	(1.404)
Farmer	-0.581	-1.171	-1.335	0.256
	(1.020)	(1.316)	(1.296)	(1.484)
Constant	10.59***	9.740***	16.99***	10.45***
	(3.460)	(3.525)	(5.014)	(3.838)
Observations	738	738	730	730
Log-likelihood	-1238.1	-1244.2	-1219.4	-1224.2

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.
Standard errors are in parentheses.

indicating lesser concavity of the utility function and lower aversion to risk.[70]

The results in Columns (1)-(4) also indicate that household wealth has a negative and statistically significant impact on loss aversion (λ). The coefficient estimates of both wealth variables are negative and significant at the 5% level in the non-IV specifications, while total livestock units and total land owned are negative and significant at the 10% and 1% levels, respectively, in the two-step specifications. Furthermore, the two-step coefficient estimates imply that the impact of wealth on loss aversion is relatively large in magnitude – a one unit increase in total livestock units causes a fall in λ of 0.75, while a one hectare increase in total land owned causes a decrease in λ of approximately 1.9.

Therefore, the results indicate that individuals from richer households are less risk averse (lower α) and less loss averse (lower λ). These results are consistent, to some extent, with those of Tanaka et al. (2010) – while they find no statistically significant relationship between household income and risk aversion (represented by the curvature parameter α) or loss aversion (represented by the parameter λ), they do find that mean village income is negatively related to both risk aversion (α) and loss aversion (λ). However, it is important to note that their IVs for household wealth are considerably weaker than those used in this study (in terms of explanatory power in the first stage regression, as indicated by the F-statistic of excluded instruments in the first stage regression); additionally, they find evidence of the endogeneity of wealth with respect to α but not λ, and thus only conduct IV estimations for the curvature parameter but not the loss aversion parameter. On the other hand, I implement IV estimations for both parameters jointly using the two-step maximum likelihood estimator, as the wealth and background risk variables are expected to be endogenously determined by both

[70] Harrison and Rutström (2009) also note that the α parameter indicates the level of risk aversion, with greater values implying greater risk aversion. Strictly speaking, in the two-part power utility specification, lower values of α imply lesser concavity of the utility function over gains (relative to the reference point). However, the reference point in this case (17 birr) is at the lower end of the spectrum of lottery prizes in the insurance problems (which range from 0 birr to 75 birr) – as a result, there are relatively few outcomes that are below the reference point (and thus evaluated as losses), and interpreting lower values of α as representing lower risk aversion (lesser concavity) is a suitable approximation.

risk and loss aversion.

Additionally, the results obtained in this section are in line with those of Tanaka and Munro (2012), who find that household wealth is negatively related to loss aversion among a sample of individuals in rural Uganda. Similarly, Liu (2008), analyzing experimental data from rural China, find some evidence of a negative relationship between wealth and the curvature parameter as well as the loss aversion parameter. These studies, however, do not account for the possible endogeneity of wealth when analyzing the determinants of risk preferences.

The large magnitude of the estimated negative impact of wealth on loss aversion in the non-EUT model highlights the possibility that, as noted by Tanaka et al. (2010) and Tanaka and Munro (2012), loss aversion may be a more important characterization of the behaviour of poor villagers in this region than risk aversion, and may play a larger role in shaping individual preferences and driving decision-making. This would not be surprising, given that the poorer individuals in rural Ethiopia live close to subsistence level, implying that any losses have a serious adverse impact on welfare (Dercon 2004). Therefore, given the importance of loss aversion in decision-making under risk, it is crucial for studies to consider non-EUT decision models and establish the determinants of loss aversion, in addition to risk aversion.

Once again, the results in Table 8 show that background risk does not have a statistically significantly effect on risk preferences, and the coefficient estimates of background risk are extremely close to zero. Additionally, in line with the EUT estimations, the results indicate that quantitative literacy has a negative impact on α – the coefficient estimate of quantitative literacy is negative and statistically significant at the 1% level in all four specifications, implying that lower quantitative literacy is associated with greater risk aversion. The estimates indicate that correctly answering one additional quantitative literacy question (out of six) leads to a decrease in α of around 0.5.

Further, the coefficient estimates of formal schooling in Columns (2) and (3) indicate that an increase in formal schooling by one year lowers λ by 0.2; these estimates are statistically significant at the 10% level.[71] Liu (2008), analyzing a similar two-part power utility function, also find that education has a negative and statistically significant relationship with loss aversion, while its relationship with risk aversion is insignificant. Additionally, the results of Tanaka et al. (2010) and Tanaka and Munro (2012) indicate that education has a negative impact on α and λ, respectively.[72,73]

An interesting result from the EUT estimations in Section 4.2 and the non-EUT estimations in this section is that gender does not have a statistically significant direct impact on risk preferences. In all the estimations of the determinants of risk preferences, the coefficient estimate of gender is statistically insignificantly different from 0. Wik et al. (2004) note that in many economies of Sub-Saharan Africa, women have greater responsibilities for providing food and caring for children, and thus are expected to be more risk averse than men – they find evidence of this using data from Northern Zambia. Additionally, Eckel and Grossman (2008) and Charness and Gneezy (2012), using developed country samples, find that women are more averse to risk than men, in line with most of the psychological and economic literature on the subject. However, Tanaka et al. (2010) find no relationship between gender and risk preferences among a sample of inhabitants of rural Vietnam, and note that earlier findings of a significant relationship in experimental studies may be due to confounds with other variables that are correlated with gender – such as education – which can be controlled for using survey data.

In order to test the robustness of the results obtained in the estimations of the

[71] Guiso and Paiella (2008) note that enrollment in higher education can depend, to some extent, on an individual's attitude towards risk. However, this is not expected to be a major issue when analyzing risk preferences in the context of rural Ethiopia, since in the sample of ERHS households, only 2% of survey respondents obtained formal schooling beyond grade nine.

[72] Indeed, Halek and Eisenhauer (2001) list numerous studies that find that education increases the propensity to take risk.

[73] It is important to note that while the estimates of the constant term for the loss aversion parameter (Table 8) are quite large in magnitude, they are consistent with those obtained by Tanaka and Munro (2012) when analyzing the determinants of loss aversion using a similar number of explanatory variables.

determinants of risk preferences, I outline and report the results of robustness checks in the following section.

5 ROBUSTNESS CHECKS

As mentioned in Section 3.1, if risk preferences drive significant internal migration of households in rural Ethiopia, doubt could be cast on the validity (exogeneity) of the instrumental variables used in this chapter. Therefore, I test the robustness of the IV results obtained in this chapter by re-estimating the models involving IVs (two-step interval regression, LIML and two-step structural maximum likelihood) after restricting the sample to those respondents who are *a priori* expected to be less likely to migrate – farmers and wealthier individuals.

Deininger and Jin (2006) and Dercon and Ayalew (2007) observe that the transferability of land rights in rural Ethiopia is severely restricted, and in most cases individuals and households residing in a particular village cannot obtain land rights in other villages. This restricts migration, particularly for farmers, who will be extremely unlikely to migrate as they cannot obtain land (or land rights) to farm in other regions (Ezra and Kiros 2001, Deininger and Jin 2006). Therefore, I conduct the IV estimations, both for the experiment and ERHS data (and for the EUT and non-EUT models), using a restricted sample of those subjects who report their primary occupation to be agricultural production. Ezra and Kiros (2001) also observe that wealthier households (primarily those that possess greater land or land rights), are less likely to migrate than poorer households due to the restricted transferability of land rights. Therefore, I also re-estimate the IV models after omitting the poorest 10% of households as measured by total land owned and then by total livestock units – these poor households are expected to be the most likely to migrate, since they would not be relinquishing any significant land rights or livestock wealth by migrating.

Using these restricted samples of households that are expected to be less likely to migrate, I obtain estimates that are similar to those obtained with the full sample, for all the IV specifications. This indicates that the IV results reported in this chapter are not driven by migration, providing some evidence that the validity of the IVs is not jeopardized by potential migration due to risk preferences, and confirming the robustness of the reported results.[74]

This is not surprising, given that voluntary migration within rural Ethiopia is mainly due to marriage (when women move to live with their husbands after marriage), and migration due to economic factors and risk preferences is very infrequent; further, voluntary migration itself is quite rare, since most of the migration in this region is forced (Ezra and Kiros 2001, Deininger and Jin 2006, Fransen and Kuschminder 2009). Indeed, Dercon and Hoddinott (2004) find a very low rate of migration among ERHS households.

In addition, given the possible endogeneity associated with household size, I conduct all the estimations in this study evaluating the determinants of risk preferences (those in Sections 4.2 and 4.3) excluding household size as an explanatory variable (as done by Tanaka et al. 2010).[75] In all cases (for both the experiment and ERHS data), the results obtained – particularly those relating to the impact of wealth and background risk – are similar to the results obtained when household size is included as an explanatory variable. This provides some indication that the possible endogeneity associated with household size is not

[74] However, it is important to acknowledge that reducing the sample size increases the likelihood of obtaining estimates that are not statistically significant.

[75] Although most studies analyzing the determinants of risk preferences in low-income regions include household or family size as an explanatory variable (for example, Yesuf and Bluffstone 2009, Tanaka and Munro 2012), household size may have a wealth effect on risk preferences in these settings – larger households represent a potentially greater labour force, and thus could generate more income and accumulate more wealth (Wik et al. 2004). Further, Wik et al. (2004) also note that larger households have greater opportunities for diversification, insurance and risk-coping. Thus, household size could affect risk preferences through its potential impact on household wealth and background risk, confounding the estimated effects of these variables on risk preferences. Furthermore, since wealth and background risk are expected to be endogenous in the context of risk preferences, this also implies that household size may be endogenous due to reverse causality. For example, more risk averse individuals may prefer to be a part of larger households, as larger households are likely to be better able to cope with risk; on the other hand, less risk averse individuals could also prefer to be a part of larger households (which probably have more wealth due to the greater potential workforce), as they would then have access to more capital to undertake high-risk (and high-return) investment and production strategies.

driving the results reported in this chapter.

Lastly, all the EUT specifications in Section 4.2 were also estimated using the basic and two-step IV versions of the structural maximum likelihood estimator (which are used for the non-EUT estimations in Section 4.3). Once again, the estimates are similar to those reported in Section 4.2, which are obtained using the basic and IV interval regression models, indicating that the reported results are not just a manifestation of the interval estimators used.

6 Conclusion

Inhabitants of poor rural economies face considerable risk, and have relatively few formal options at their disposal for coping with this substantial risk. Given the prevalence of uninsured risk, risk preferences play a significant role in establishing the investment, production and risk-coping strategies of rural inhabitants. Thus, analyzing risk preferences, and their determinants, in low-income settings is crucial, since they have a significant impact on the wealth and welfare of households.

In this chapter, I evaluate the determinants of risk preferences in rural Ethiopia and assess the extent of asset integration in experimental decisions, using combined experimental and survey data. I focus on estimating the effect of two economic factors – wealth and background risk – on risk attitudes. Establishing the direction and magnitude of the impact of these economic circumstances on risk preferences is essential for a better understanding of decision-making under risk, and has crucial policy implications (Banerjee 2000, Falco 2012). I account for the possible endogeneity associated with wealth and background risk by using instrumental variables in combination with two-step and Limited Information Maximum Likelihood interval estimators (described by Bettin and Lucchetti 2010), and a two-step version of the structural maximum likelihood estimator detailed in the work of Harrison and Rutström (2008b). To my knowledge, this analysis represents the first application of these estimators to the evaluation of the determinants of risk preferences, and also one of the few instances where the endogeneity of both wealth and background risk is accounted for by the use of instrumental variables. Further, I use both EUT and non-EUT decision models to analyze choices from Binswanger-style lotteries and decision problems involving insurance purchase, respectively; on the other hand, most studies on the subject consider only the single-parameter EUT-CRRA framework.

I find some evidence that household wealth negatively affects both risk aversion and loss aversion. This implies that richer households are less averse to both risk and losses than poorer households, and is in line with the results of recent experimental studies conducted in developing countries (for example, Yesuf and Bluffstone 2009, Tanaka et al. 2010, Tanaka and Munro 2012).[76] This finding, combined with evidence from numerous studies on the path dependence of wealth accumulation, indicates that risk preferences could have a crucial impact on long-run poverty – poorer households are more risk and loss averse and thus may be more likely to undertake low-risk, low-return investment and production strategies, possibly hampering wealth accumulation and leading to persistent poverty; Mosley and Verschoor (2005) refer to this as the "vicious circle of poverty" in which risk attitudes play a crucial role. Therefore, policies that promote rural development and increase household wealth could have a significant impact on long-term (in addition to short-term) poverty and welfare, through mitigating aversion to risk (as well as losses) and increasing involvement in high-risk, high-return production and investment strategies (Yesuf and Bluffstone 2009).

Using the instrumental variable analysis in combination with both the EUT and non-EUT decision models, I find no evidence that independent background risk impacts the risk preferences of individuals – the coefficient estimates

[76] Mosley and Verschoor (2005) find that the state of mind brought about by chronic poverty, measured by an index of perceived vulnerability, causes greater risk aversion. They note that the psychological repercussions of chronic material hardship lead to an outlook on life referred to as an "external locus of control", which hinders risk-taking.

of background risk are both statistically insignificant and extremely small in magnitude. Thus, I do not find support for Gollier and Pratt's (1996) risk vulnerability hypothesis, in line with the work of Alessie et al. (2002) and Lusk and Coble (2008).

Further, I find evidence of narrow framing, as opposed to asset integration, suggesting that participants isolate decisions in the experimental lotteries from other decisions in their lives, and make decisions by considering only the prospective gains and losses associated with the current decision, independent of outside wealth. This is consistent with the results of many experimental studies conducted in both developed and developing countries (Heinemann 2008). According to Lusk and Coble (2008), this provides some evidence that individuals assess each risky choice in isolation, which could cause them to disregard, to a certain extent, background risks when making decisions in experimental lotteries – this, in turn, would lead to a lower estimated effect of background risk on risk preferences than predicted by the theoretical models of Gollier and Pratt (1996) and Quiggin (2003). Alternatively, the insignificant estimates of background risk could arise because part of its effect on risk attitudes is captured by the wealth variable – as noted by Yesuf and Bluffstone (2009), wealth accumulation often substitutes for missing markets in developing economies and alleviates credit and insurance constraints, implying that wealthier households are probably better at insulating themselves from shocks and thus are exposed to lesser background risk.

Additionally, most studies that do find a significant positive relationship between background risk and risk aversion focus on experimentally-induced background risk (for example, Harrison et al. 2007, Lee 2008), or use data from developed countries (for example, Heaton and Lucas 2000, Guiso and Paiella 2008). Relatively little attention has been given in the experimental literature to the impact of non-experimental background risk on risk preferences exhibited in laboratory experiments conducted in low-income rural regions. Further, Lusk and Coble (2008) note that experimental subjects are likely to bring numerous background risks with them into an experiment, and thus analyzing the impact of these non-experimental ("field") background risks on risk attitudes may be more important, as the effect of these risks might dominate the impact of experimentally-induced background risk. Jointly assessing the impact of experimental and non-experimental background risks on risk preferences in a single experiment could shed light on the relative importance of the two in determining risk preferences, and represents an interesting avenue for further research.

While the results obtained in this study indicate that gender does not have a statistically significant impact on risk attitudes, I do find some evidence that formal education and quantitative literacy affect risk preferences – the estimated impact of these variables on loss aversion and on risk aversion is negative and statistically significant in the non-EUT estimations, using data from the insurance decision problems. This indicates that policies which improve schooling and quantitative literacy could increase income (and wealth) not only through increases in human capital and productivity, but also through lowering aversion to risk and thus raising the propensity to undertake risky production and investment strategies with higher average returns (Tanaka and Munro 2012).

As noted by Harrison et al. (2007), laboratory experiments conducted in the field with local subjects – such as the one analyzed in this chapter – represent "a useful middle ground between the tight controls of the laboratory and the vagaries of completely uncontrolled field data". The greater degree of control associated with the laboratory environment and related experimental methods enables the observation of decision-making under risk in isolation from other factors that could influence behaviour, and thus is crucial for the evaluation of the determinants of risk preferences; further, conducting the experimental decision problems with poor subjects in rural Ethiopia – who are routinely faced with risk – increases the external validity of the experimental results (Harrison and List 2004). Additionally, the combination of experiment and survey data from the ERHS provides fertile ground to analyze the asset integration hypothesis, and to estimate the sign and magnitude of the impact of an extensive set of demographic and socioeconomic factors on risk preferences. Thus, such combined datasets containing experimental and detailed survey data represent an important tool for future research on decision-making under risk, and facilitate a more comprehensive characterization of risk preferences and their determinants.

REFERENCES

Alessie, RJM, S. Hochgürtel, and A.H.O. Soest, “Household portfolios in the Netherlands,” in L. Guiso, M. Haliassos, and T. Jappelli, eds., *Household Portfolios*, MIT Press, 2002, pp. 341–388.

Andersen, S., G.W. Harrison, and E.E. Rutström, “Choice behavior, asset integration and natural reference points,” *Department of Economics Working Papers*, 2006.

——, ——, M.I. Lau, and E.E. Rutström, “Dynamic choice behavior in a natural experiment,” *University of Central Florida Department of Economics Working Paper*, 2006, pp. 06–10.

——, ——, ——, and ——, “Risk aversion in game shows,” 2008.

Arrow, K.J., *Essays in the Theory of Risk Bearing*, Chicago: Markham, 1971.

Ashraf, N., D. Karlan, and W. Yin, “Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines,” *Quarterly Journal of Economics*, 2006, *121* (2), 635–672.

Attanasio, Orazio, Abigail Barr, Juan Camilo Cardenas, Garance Genicot, and Costas Meghir, “Risk Pooling, Risk Preferences, and Social Networks,” *American Economic Journal: Applied Economics*, 2012, *4* (2), 134–167.

Ayalew, D.A., S. Dercon, and M. Gautam, “Property rights in a very poor country: tenure insecurity and investment in Ethiopia.,” *Global Poverty Research Group Working Paper*, 2005.

Banerjee, A.V., “The Two Poverties,” *Nordic Journal of Political Economy*, 2000, *26* (2), 129–141. Reprinted in Dercon, Stefan (ed), “Insurance Against Poverty”, Oxford University Press, 2004.

—— and A.F. Newman, “Occupational Choice and the Process of Development,” *The Journal of Political Economy*, 1993, *101* (2), 274–298.

Barr, Abigail and Truman G Packard, “Seeking solutions to Vulnerability in Old Age: Preferences, constraints, and alternatives for coverage under PeruÕs pension system,” *The Centre for the Study of African Economies Working Paper Series*, 2005, p. 239.

Barrett, C. and J. McPeak, “Poverty traps and safety nets,” *Poverty, Inequality and Development*, 2006, pp. 131–154.

Bettin, G. and R. Lucchetti, “Interval regression models with endogenous explanatory variables,” *Empirical Economics*, 2010, pp. 1–24.

Binswanger, H.P., “Attitudes toward risk: Experimental measurement in rural India,” *American Journal of Agricultural Economics*, 1980, *62* (3), 395–407.

——, “Attitudes toward risk: Theoretical implications of an experiment in rural India,” *Economic Journal*, 1981, *91*, 867–890.

Botelho, A., G.W. Harrison, L. Pinto, E.E. Rutström, and P. Veiga, “Discounting in developing countries: a pilot experiment in Timor-Leste,” *NIMA Working Papers 31*, 2005.

Burks, Stephen V, Jeffrey P Carpenter, Lorenz Goette, and Aldo Rustichini, "Cognitive skills affect economic preferences, strategic behavior, and job attachment," *Proceedings of the National Academy of Sciences*, 2009, *106* (19), 7745–7750.

Camerer, C.F. and T.H. Ho, "Violations of the betweenness axiom and nonlinearity in probability," *Journal of Risk and Uncertainty*, 1994, *8* (2), 167–196.

Charness, G. and U. Gneezy, "Strong Evidence for Gender Differences in Risk Taking," *Journal of Economic Behavior & Organization*, 2012, *83* (1), 50–58.

Charness, Gary, Uri Gneezy, and Alex Imas, "Experiential methods: Eliciting risk preferences," *Journal of Economic Behavior & Organization*, 2013.

Clarke, K.A., "Nonparametric model discrimination in international relations," *Journal of Conflict Resolution*, 2003, *47* (1), 72.

Cole, Shawn A., Xavier Giné, Jeremy B. Tobacman, Petia B. Topalova, Robert M. Townsend, and James I. Vickery, "Barriers to Household Risk Management: Evidence from India," Working Paper 09-116, Harvard Business School 2009.

Cox, J.C. and V. Sadiraj, "Implications of small-and large-stakes risk aversion for decision theory," *Working paper prepared for workshop on Measuring Risk and Time Preferences by the Centre for Economic and Business Research in Copenhagen*, 2004.

____ **and** ____ , "Small-and large-stakes risk aversion: Implications of concavity calibration for decision theory," *Games and Economic Behavior*, 2006, *56* (1), 45–60.

Dave, C., C.C. Eckel, C.A. Johnson, and C. Rojas, "Eliciting risk preferences: When is simple better?," *Journal of Risk and Uncertainty*, 2010, pp. 1–25.

Davidson, R. and J.G. MacKinnon, "Estimation and inference in econometrics," *OUP Catalogue*, 1993.

Deininger, K. and S. Jin, "Tenure security and land-related investment: Evidence from Ethiopia," *European Economic Review*, 2006, *50* (5), 1245–1277.

Dercon, S., "Growth and shocks: evidence from rural Ethiopia," *Journal of Development Economics*, 2004, *74* (2), 309–329.

____ **and D. Ayalew**, "Land Rights, Power and Trees in Rural Ethiopia," *CSAE Working Paper*, 2007.

____ **and J. Hoddinott**, "The Ethiopian rural household surveys: Introduction," *International Food Policy Research Institute, Washington, DC*, 2004.

____ **and P. Krishnan**, "Changes in poverty in rural Ethiopia 1989-1995: measurement, robustness tests and decomposition," *CSAE Working Paper Series*, 1998, *98.7*.

____ **and** ____ , "In Sickness and in Health: Risk Sharing within Households in Rural Ethiopia," *The Journal of Political Economy*, 2000, *108* (4), 688–727.

Dercon, Stefan, Ruth Vargas Hill, Ingo Outes-Leon, A Bayrou, D Clarke, and A Seyoum-Taffesse, "Offering rainfall insurance to informal insurance groups: evidence from a field experiment in Ethiopia," Technical Report, Mimeo 2011.

Diamond, D.W., "Financial intermediation and delegated monitoring," *The Review of Economic Studies*, 1984, *51* (3), 393–414.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde, "Are Risk Aversion and Impatience Related to Cognitive Ability?," *American Economic Review*, 2010, *100* (3).

Eckel, C.C. and P.J. Grossman, "Differences in the economic decisions of men and women: Experimental evidence," *Handbook of experimental economics results*, 2008, *1*, 509–519.

Engle-Warnick, J., J. Escobal, and S. Laszlo, "Ambiguity Aversion and Portfolio Choice in Small-Scale Peruvian Farming," *The B.E. Journal of Economic Analysis and Policy*, 2011, *11* (1).

Ezra, M. and G.E. Kiros, "Rural Out-migration in the Drought Prone Areas of Ethiopia: A Multilevel Analysis1," *International Migration Review*, 2001, *35* (3), 749–771.

Fafchamps, M., *Rural Poverty, Risk and Development*, Edward Elgar Publishing, 2003.

____ **and A. Quisumbing**, "Assets at marriage in rural Ethiopia," *Journal of Development Economics*, 2005, *77* (1), 1–25.

____ **, C. Udry, and K. Czukas**, "Drought and saving in West Africa: are livestock a buffer stock?," *Journal of Development economics*, 1998, *55* (2), 273–305.

Fafchamps, Marcel, "Vulnerability, Risk Management, and Agricultural Development," *CSAE Working Paper Series*, 2009.

____ **, Bereket Kebede, and Daniel John Zizzo**, "Keep Up With the Winners: Experimental Evidence on Risk Taking, Asset Integration, and Peer Effects," 2013.

Falco, P., "Does risk matter for occupational choices? Experimental Evidence from an African labour market," *Manuscript submitted for publication*, 2012.

Foster, A.D. and M.R. Rosenzweig, "Learning by doing and learning from others: Human capital and technical change in agriculture," *Journal of political Economy*, 1995, pp. 1176–1209.

Fransen, S. and K. Kuschminder, "Migration in Ethiopia: History, Current Trends and Future Prospects," *Paper Series: Migration and Development Country Profile, Maastricht Graduate School of Governance (MGSoG)*, 2009.

Gollier, C. and J.W. Pratt, "Risk vulnerability and the tempering effect of background risk," *Econometrica: Journal of the Econometric Society*, 1996, pp. 1109–1123.

Guiso, L. and M. Paiella, "Risk aversion, wealth, and background risk," *Journal of the European Economic Association*, 2008, *6* (6), 1109–1150.

Halek, M. and J.G. Eisenhauer, "Demography of risk aversion," *Journal of Risk and Insurance*, 2001, pp. 1–24.

Harrison, G.W. and E.E. Rutström, "Experimental evidence on the existence of hypothetical bias in value elicitation methods," *Handbook of experimental economics results*, 2008, *1*, 752–767.

____ **and E.E. Rutström**, "Risk aversion in the laboratory," in J. C. Cox and G. W. Harrison, eds., *Research in Experimental Economics*, Vol. 12, Emerald Group Publishing Limited, 2008, pp. 41–196.

____ **and** ____ , "Expected utility theory and prospect theory: One wedding and a decent funeral," *Experimental Economics*, 2009, *12* (2), 133–158.

____ **and J.A. List**, "Field experiments," *Journal of Economic Literature*, 2004, *42* (4), 1009–1055.

____ **,** ____ **, and C. Towe**, "Naturally occurring preferences and exogenous laboratory experiments: A case study of risk aversion," *Econometrica*, 2007, *75* (2), 433–458.

____ **, M.I. Lau, and E.E. Rutström**, "Theory, experimental design and econometrics are complementary (and so are lab and field experiments)," *The Methods of Modern Experimental Economics*, 2011.

____ **, S.J. Humphrey, and A. Verschoor**, "Choice under Uncertainty: Evidence from Ethiopia, India and Uganda," *The Economic Journal*, 2010, *120* (543), 80–104.

Heaton, J. and D. Lucas, "Portfolio choice in the presence of background risk," *The Economic Journal*, 2000, *110* (460), 1–26.

Heinemann, F., "Measuring risk aversion and the wealth effect," *Research in Experimental Economics*, 2008, *12*, 293–313.

Hey, J.D., "Experimental investigations of errors in decision making under risk," *European Economic Review*, 1995, *39* (3-4), 633–640.

Hill, R.V., J. Hoddinott, and N. Kumar, "Adoption of weather index insurance: Learning from willingness to pay among a panel of households in rural Ethiopia," *IFPRI discussion paper 01088*, 2011.

Holt, C.A. and S.K. Laury, "Risk aversion and incentive effects," *American Economic Review*, 2002, *92* (5), 1644–1655.

Hong, H. and E. Tamer, "Inference in censored models with endogenous regressors," *Econometrica*, 2003, *71* (3), 905–932.

Jalan, J. and M. Ravallion, "Transient poverty in postreform rural China," *Journal of Comparative Economics*, 1998, *26* (2), 338–357.

Kamanou, G. and J. Morduch, "8. Measuring Vulnerability to Poverty," *Insurance Against Poverty*, 2004, *1* (9), 155–176.

Lee, J., "The effect of the background risk in a simple chance improving decision model," *Journal of Risk and Uncertainty*, 2008, *36* (1), 19–41.

Liu, Elaine M, "Time to change what to sow: Risk preferences and technology adoption decisions of cotton farmers in China," *Review of Economics and Statistics*, 2008.

Lusk, J.L. and K.H. Coble, "Risk aversion in the presence of background risk: Evidence from an economic experiment," *Research in Experimental Economics*, 2008, *12*, 315–340.

Lybbert, Travis J, Francisco B Galarza, John McPeak, Christopher B Barrett, Stephen R Boucher, Michael R Carter, Sommarat Chantarat, Aziz Fadlaoui, and Andrew Mude, "Dynamic Field Experiments in Development Economics: Risk Valuation in Morocco, Kenya, and Peru," *Agricultural and Resource Economics Review*, 2010, *39*, 1–17.

Manski, C.F. and E. Tamer, "Inference on regressions with interval data on a regressor or outcome," *Econometrica*, 2002, *70* (2), 519–546.

Masson, A. and L. Arrondel, "Gestion du risque et comportements patrimoniaux," *Économie et Statistique*, 1996, *296* (1), 63–89.

Moscardi, E. and A. de Janvry, "Attitudes toward risk among peasants: an econometric approach," *American Journal of Agricultural Economics*, 1977, *59* (4), 710–716.

Mosley, P. and A. Verschoor, "Risk attitudes and the 'vicious circle of poverty'," *The European Journal of Development Research*, 2005, *17* (1), 59–88.

Munshi, K. and M. Rosenzweig, "Why is mobility in India so low? Social insurance, inequality, and growth," *National Bureau of Economic Research Working Paper*, 2009.

Nielsen, U., *Poverty and attitudes towards time and risk: Experimental evidence from Madagascar*, Royal Veterinary and Agricultural Univ., 2001.

Oehlert, G.W., "A note on the delta method," *American Statistician*, 1992, *46* (1), 27–29.

Porter, C., "Examining the impact of idiosyncratic and covariate shocks on Ethiopian households' consumption and income sources," Technical Report 2011.

Powell, M. and D. Ansic, "Gender differences in risk behaviour in financial decision-making: An experimental analysis," *Journal of Economic Psychology*, 1997, *18* (6), 605–628.

Quiggin, J., "Background risk in generalized expected utility theory," *Economic Theory*, 2003, *22* (3), 607–611.

Rabin, M., "Risk aversion and expected-utility theory: A calibration theorem," *Econometrica*, 2000, *68* (5), 1281–1292.

Rivers, D. and Q.H. Vuong, "Limited information estimators and exogeneity tests for simultaneous probit models," *Journal of Econometrics*, 1988, *39* (3), 347–366.

Rosenzweig, M.R. and H.P. Binswanger, "Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments," *The Economic Journal*, 1993, *103* (416), 56–78.

Schechter, L., "Risk aversion and expected-utility theory: A calibration exercise," *Journal of Risk and Uncertainty*, 2007, *35* (1), 67–76.

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

Tanaka, Y. and A. Munro, "Regional Variation in Risk and Time Preferences: Evidence from a Large-Scale Field Experiment in Rural Uganda," *National Graduate Institute for Policy Studies Working Paper, Tokyo*, 2012.

Tversky, A. and D. Kahneman, "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and uncertainty*, 1992, *5* (4), 297–323.

Vuong, Q.H., "Likelihood ratio tests for model selection and non-nested hypotheses," *Econometrica: Journal of the Econometric Society*, 1989, *57* (2), 307–333.

Walker, T.S. and J.G. Ryan, *Village and household economics in India's semi-arid tropics*, Johns Hopkins University Press, 1990.

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

Wooldridge, J.M., *Econometric analysis of cross section and panel data*, The MIT press, 2002.

Yesuf, M. and R.A. Bluffstone, "Poverty, risk aversion, and path dependence in low-income countries: Experimental evidence from Ethiopia," *American Journal of Agricultural Economics*, 2009, *91* (4), 1022–1037.

Zyphur, Michael J, Jayanth Narayanan, Richard D Arvey, and Gordon J Alexander, "The genetics of economic risk preferences," *Journal of Behavioral Decision Making*, 2009, *22* (4), 367–377.

APPENDIX A: QUANTITATIVE LITERACY

Table 9. *Questions assessing quantitative literacy*

Variables	No. of Observations	Mean	Standard Deviation
Experiment Participants			
Yellow or blue token draw more likely?	378	0.884	0.321
$5 + 3 =?$	378	0.862	0.345
$3 \times 7 =?$	378	0.545	0.498
$\frac{1}{10}$th of $300 =?$	378	0.300	0.458
5% of $200 =?$	378	0.0132	0.114
Riskier to plant one crop or multiple crops?	378	0.852	0.356
Round 7 ERHS Respondents			
Lion or crown more likely in a coin flip?	1343	0.456	0.498
White or red ball draw more likely from a bag with 1 white and 4 red balls?	1347	0.828	0.378
$4 + 3 =?$	1347	0.915	0.278
$3 \times 6 =?$	1347	0.709	0.454
$\frac{1}{10}$th of $400 =?$	1346	0.395	0.489
15% of $200 =?$	1346	0.282	0.450

APPENDIX B: FIRST STAGE RESULTS FOR TWO-STEP INTERVAL REGRESSION

Table 10. *First stage OLS results for two-step interval regression: Experiment Binswanger-style lottery*

Variables	(1) Total livestock units	(2) Total land owned	(3) Std. dev. of consumption
Inherited land	0.139	0.165*	41.07*
	(0.159)	(0.0952)	(21.70)
Distance to town * Gender	0.107***	-0.0281	-2.209
	(0.0299)	(0.0209)	(4.002)
Village size * Gender	-0.00104*	-0.00244***	0.276***
	(0.000595)	(0.000506)	(.0931)
Std. dev. of rainfall * Gender	-0.00926***	-0.00545	-1.808***
	(0.00132)	(0.00805)	(0.264)
Observations	367	367	365
R-squared	0.308	0.0642	0.255
F-statistic of excluded instruments	32.68	15.68	13.44
p-value joint significance F-test of excluded IVs	0.000	0.000	0.000

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Standard errors are in parentheses.

Only identifying regressors (excluded instruments) reported.

Table 11. *First stage OLS results for two-step interval regression: ERHS Binswanger-style lottery*

Variables	(1) Total livestock units	(2) Total land owned	(3) Std. dev. of consumption
Inherited land	-0.0646	0.413***	25.98*
	(0.0931)	(0.0639)	(15.07)
Distance to town * Gender	-0.0446***	-0.0316*	3.046
	(0.0109)	(0.0174)	(2.144)
Village size * Gender	0.000223***	-0.0000457	0.0282***
	(0.0000349)	(0.0000778)	(0.00810)
Std. dev. of rainfall * Gender	-0.000130	0.00313	0.597***
	(0.00126)	(0.00267)	(0.156)
Observations	1255	1255	1255
R-squared	0.428	0.164	0.341
F-statistic of excluded instruments	15.10	12.18	7.94
p-value joint significance F-test of excluded IVs	0.000	0.000	0.000

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$
Standard errors are in parentheses.
Only identifying regressors (excluded instruments) reported.

ARE RISK PREFERENCES REVEALED BY EXPERIMENTAL CHOICE? EVIDENCE FROM BRAZIL

GAUTAM KALANI*

Abstract

Do experimental subjects understand decision problems that use the relatively complex Multiple Price List (MPL) elicitation procedure? To answer this important methodological question, and to determine which decision models best describe the decision-making under risk of inhabitants, I consider a MPL experiment from Brazil. I find that the MPL decision problems, in which participants answer a series of 20 binary questions, enable a finer characterization of risk preferences – and the estimation of a greater range of decision models – as compared to Ordered Lottery Selection problems. However, I find that a significant fraction of choice patterns in the decision problems are intransitive, and obtain evidence of substantial noise in decision-making. Further, the evidence indicates that poor comprehension is the most likely cause of the large fraction of inconsistent choices observed, rather than indifference or random errors. I also find that an alternative decision model – which does not assume transitivity in choice and is not based on well-defined preferences over risk – better fits the data than any of the standard models considered. The results indicate that subjects did not properly understand the MPL decision problems and thus observed choices do not reveal true risk preferences. Therefore, the complex MPL procedure may not be suitable for experiments conducted with poorly-educated subjects in developing country settings, and simpler procedures are likely to be preferred in these settings.

JEL codes: C91, D01, D03, D81, G22, O16.

*DPhil student: Department of Economics, Centre for the Study of African Economies (CSAE) and Hertford College, University of Oxford (**gautam.kalani@economics.ox.ac.uk**). This paper is part of the work towards my DPhil thesis, which is supervised by Professor Marcel Fafchamps; the work would not have been possible without his very generous assistance. Additionally, I am grateful to Daniel Osgood, Maria Velez, Alexander Pfaff, Simon Quinn and Daniel Clarke, who offered useful suggestions throughout the project. Thanks also to participants at the Gorman Student Research Workshop in Oxford University, where a version of this paper was presented.

1 INTRODUCTION

In Ordered Lottery Selection decision problems (popularized by Binswanger 1980), such as those comprising the Ethiopian experiment, participants choose one out of a number of options (generally four to six). On the other hand, in decision problems using the Multiple Price List (MPL) elicitation procedure popularized by Holt and Laury (2002), participants choose their preferred option from a pair of prospects, for a series of such pairs (generally 10-30). Each choice in an Ordered Lottery Selection decision problem therefore provides more information on risk preferences than a choice among a pair of prospects in a Multiple Price List (MPL) decision problem. However, since a MPL decision problem involves a relatively large number of choices, it enables the classification of subjects into a larger number of risk preference categories – and thus facilitates a more precise calibration of risk attitudes – than an Ordered Lottery Selection decision problem, which involves only a single choice (Dave et al. 2010).[1]

This greater precision, combined with the larger number of total choices in the dataset, associated with the MPL procedure makes it more suitable for obtaining accurate point estimates of preference parameters using econometric methods (Andreoni and Sprenger 2012, Charness et al. 2013). Therefore, the results of Chapter 1, in which I estimate and analyze various decision models using experimental data from rural Ethiopia, indicate that certain choice models cannot be estimated and certain preference parameters cannot be identified using experimental data from decision problems of the type developed by Binswanger (1980). However, decision problems using the complex MPL elicitation technique are also cognitively more challenging (that is, they place a higher "cognitive load" on subjects) since they involve more choices (thus requiring more complex mental calculations and greater mathematical sophistication in

[1] Following the nomenclature used in Chapter 1, I refer to any decision problem that involves multiple binary choices as utilizing the Holt and Laury (2002) experimental design, and use this term interchangeably with the Multiple Price List elicitation procedure. Similarly, I refer to any decision problem involving a single choice between more than two options as utilizing the Binswanger (1980) experimental design, and use this term interchangeably with the Ordered Lottery Selection elicitation procedure and Binswanger-style lottery.

order to make consistent choices). As a result, some experimental studies find that individuals, particularly those with low levels of education and poor mathematical skills, do not understand decision problems using this elicitation procedure and their responses are characterized by substantial noise (Dave et al. 2010, Crosetto and Filippin 2013).

In order to evaluate whether the MPL procedure enables a finer characterization of risk preferences and thus the estimation of a wider range of decision models, and to analyze whether subjects face difficulty in comprehending MPL decision problems, I use data from a lab experiment in Brazil that featured decision problems using this elicitation procedure.[2] In each decision problem of this experiment, which involved a combination of urban college students and rural agricultural producers, participants were required to answer 20 binary questions indicating whether or not they would prefer to buy a particular insurance contract at different prices. I estimate all the major decision models analyzed in Chapter 1 using the structural maximum likelihood estimation methodology (described by Harrison and Rutström 2008); the decision models estimated include the "standard" decision theories – expected utility theory (EUT), rank-dependent utility (RDU) and cumulative prospect theory (CPT).[3] This analysis represents an appropriate complement to that conducted in Chapter 1 – it provides an indication of which preference parameters are identified and can be accurately estimated using Multiple Price List decision problems as compared to Ordered Lottery Selection problems, thus shedding light on the relative power provided by the data from the Brazilian and Ethiopian experiments. Further, this analysis also facilitates the determination of which decision models (and associated preference parameters) best describe the behaviour under risk of individuals in the Northeast region of Brazil, which is characterized by drought, low agricultural productivity and persistent poverty (Broad et al. 2007). Understanding and characterizing the

[2] Crosetto and Filippin (2013) note that, to date, the MPL mechanism is the most popular risk elicitation procedure.

[3] These decision models are referred to as "standard" as they are the major choice models used by most experimental studies to describe the behaviour of individuals in decisions involving risk (Hey and Orme 1994, Harrison and Rutström 2008).

decision-making under risk of low-income populations is an important objective of this thesis, and that theme is prevalent in this chapter as well.

I do find that the MPL design of the Brazilian experiment enables the estimation of a greater range of decision models as compared to the data from the Ethiopian experiment, thus enabling a more comprehensive analysis of risk attitudes. In addition to fitting the standard decision models listed above to the Brazilian experimental data, I also analyze whether subjects' choices conform to the predictions of these models and explore the extent of stochastic errors (or noise) in decision-making, using the Luce and Fechner error models. In line with the observations of Dave et al. (2010), I find that if true preferences are assumed to follow any of the standard decision models considered, then large estimates of the noise parameter are obtained in maximum likelihood estimations of these error models.

Additionally, I find that a substantial fraction (33%) of choice patterns in the MPL decision problems are intransitive – due to multiple or reverse switching – and thus are not consistent with any of the standard decision models, or indeed with any well-defined preferences over risk. Barr (2007) notes that the presence of such inconsistent choice patterns could indicate a lack of understanding of the decision problems, random errors in decision-making, or genuine indifference between the alternatives. An important aim of this chapter is to determine which of these factors is the most likely cause of the inconsistent choice patterns observed in the Brazilian experiment. This issue has not been analyzed in depth in the experimental literature; most experimental studies ignore or discard such intransitive choices assuming them to be uninformative noise. However, Jacobson and Petrie (2009) stress that these choices could provide valuable insights into the decision-making process under risk of individuals.

I find that indifference between alternatives in the decision problems is unlikely to be a cause of the inconsistent patterns observed in the data. Further, using data on participant characteristics collected in conjunction with the experiment to estimate Probit and Ordered Probit specifications, I find that participants who did not attend college are significantly more likely to have intransitive choice patterns, indicating that the observed inconsistencies could be related to low cognitive ability. Using structural maximum likelihood estimations, I also find that the noise parameter in the Luce and Fechner error models is negatively related to college education.

In order to test whether the observed inconsistencies in choice are due to random ("trembling hand") errors which cause deviations in observed choices from true transitive preferences or due to the true decision-making process of participants (which could be non-transitive), I utilize the non-parametric constant error rate analysis of Harless and Camerer (1994). To my knowledge, the constant error rate analysis has not been previously applied in this context using MPL data involving a relatively large number of choices in each decision problem, such as in the Brazilian experiment. Using this analysis, I reject the hypothesis that the inconsistent choice patterns are caused by random errors, and the results indicate that these choices are likely to be a result of the true decision-making process which is non-transitive.[4]

Further, I estimate two non-transitive, alternative models of decision-making and investigate whether they provide a better description of observed choices than the standard decision models. Using Vuong tests and mixture model specifications (as described by Harrison and Rutström 2009), I find that an alternative decision model in which participants choose a particular option in the first binary question of a decision problem with a particular fixed probability and then in subsequent questions stick to their previous choice with another fixed probability better fits the data than any of the standard decision theories. However, since this alternative model is not based on rational decision-making with well-defined risk preferences and generates intransitive choice patterns, the indication that it best represents the latent decision-making process of subjects implies that the choices of participants

[4] In other words, I reject the hypothesis that choices are generated by noisy rational choice (that is, by rational choice plus error).

in the MPL decision problems do not reveal true risk preferences (which are presumably transitive), but rather provide information on subjects' preferences about how to play the "games" in a way that is not completely determined by preferences over risk.

Therefore, the results in this chapter indicate that participants most likely use the metric or heuristic implied by this alternative model for making decisions – thus essentially making choices in a psuedo-random fashion – due to confusion and poor understanding of the decision problems; this, rather than indifference or "trembling hand" errors, is the most likely cause of the large fraction of inconsistent choices observed in the data, highlighting the issue of poor understanding associated with the MPL procedure.

This is the major methodological contribution of this chapter to the experimental economics literature, and to experimental development economics in particular; as noted by Dave et al. (2010), work on this topic is part of the methodological effort to improve experimental design and enhance the measurement of risk preferences. Similar results are obtained by Charness and Viceisza (2012), who find that a majority of participants involved in a lab experiment – which includes three different elicitation methods – conducted in rural Senegal do not properly understand the complex Holt and Laury (2002) elicitation procedure, and as a result make inconsistent choices that do not reflect true risk preferences. Dave et al. (2010), who conduct the Ordered Lottery Selection and MPL procedures with the same sample of Canadian subjects, find that participants with low mathematical ability and education levels, in particular, face much more difficulty in understanding the MPL procedure as compared to the Ordered Lottery Selection method, leading to substantially more randomness in their choices in the MPL decision problem. Thus, Crosetto and Filippin (2013) conclude that while the difference in complexity between the MPL and Ordered Lottery Selection experimental procedures may not be an issue for a standard experimental sample of university students, subjects with low levels of education find the Ordered Lottery Selection method significantly easier to understand.

This chapter therefore provides a methodological and empirical contribution, and builds on the work of Dave et al. (2010), Charness and Viceisza (2012) and Crosetto and Filippin (2013) in assessing the suitability of the MPL procedure for accurately eliciting risk preferences. However, while these studies conduct different experimental procedures with the same sample of subjects, they primarily consider expected utility theory – thus, the large estimates of the noise parameter they obtain using data from MPL problems may be because the decision-making process of participants is not accurately reflected by EUT, rather than due to randomness in choice caused by poor understanding. Therefore, assuming an incorrect decision-making process could yield misleading results and biased estimates of preference parameters (Charness et al. 2013); this highlights the importance of considering a wide range of different decision models in analyses of behaviour under risk. Additionally, the above-mentioned studies do not test between the different possible causes of the observed inconsistencies in choice, and most studies on this topic analyze data from experiments conducted in developed countries (with the exception of Charness and Viceisza 2012).

The results of studies (for example, Dave et al. 2010, Crosetto and Filippin 2013) that use a direct comparison of noise parameter estimates (in Luce and Fechner error models) to compare the extent of errors in decision-making across contexts in which the preference parameters differ could be misleading; this is because variation in the noise parameter does not solely capture differences in noise, as the noise parameter estimate depends also on the preference parameter values. Therefore, I propose a normalization procedure which enables the direct comparison of the noise parameter across contexts in which preference parameters differ – this represents another important contribution of this chapter. Comparing the normalized noise parameter estimates across the Brazilian and Ethiopian experiments analyzed in this thesis, I find that the magnitude of errors in decision-making is greater in the Brazilian experiment, likely due to poor understanding of the complex MPL procedure. Thus, the results in this chapter, in line with those of Charness and Viceisza (2012), indicate that the relatively complex MPL mechanism may not be suitable for eliciting consistent choices that accurately reflect true risk preferences in developing country settings characterized by low education levels. Instead, as noted by Charness et al. (2013), simpler procedures, such as the Ordered Lottery Selection method, may be more appropriate for experiments conducted in these settings.

Lybbert et al. (2010) note that complexity and participant understanding of decision problems are crucial factors to consider when conducting experiments in developing countries with poorly-educated subjects; however, the analysis of these issues has been given relatively little attention in the experimental development economics literature, despite the observation that tasks suitable for standard experimental samples may not be appropriate for individuals with limited formal education. The results in this chapter also highlight the importance of not implicitly assuming transitivity in experimental choice (as noted by Fishburn 1991), and thus experiments which use variations of the MPL procedure that impose consistency in choice (for example, Liu 2008, Tanaka et al. 2010) could yield misleading inferences regarding decision-making under risk.

The rest of this chapter is organized as follows. The next section outlines the empirical investigation strategy, while Section 3 provides a description of the experiment design as well as the associated data on participant characteristics. The empirical analysis is presented in Section 4 – the estimates of the standard decision models are presented in Section 4.1, the analysis of the correlates of stochastic errors and inconsistencies in choice is presented in Section 4.2, and the alternative decision models as well as the constant error rate analysis are considered in Section 4.3. Section 5 concludes.

2 OUTLINE OF EMPIRICAL INVESTIGATION

In this chapter, I estimate all the decision models described in Section 2 of Chapter 1. I first estimate EUT constant relative risk aversion (EUT-CRRA) and EUT expo-power (EUT-EP) models. While EUT is still considered by many to be the dominant theory of decision-making under risk in economics, numerous experimental studies both in developed and developing countries have found strong evidence of loss aversion and probability weighting, with decisions conforming to rank-dependent utility (RDU) and cumulative prospect theory (CPT) – thus, RDU and CPT are generally considered to be the strongest contenders to EUT as theories of decision-making under risk (Hey and Orme 1994, Humphrey and Verschoor 2004a, Tanaka et al. 2010). Therefore, in addition to the EUT specifications, I also estimate a two-part power utility (PU) model (as described by Tversky and Kahneman 1992), Quiggin's (1982) RDU – combined with a CRRA and an EP utility function – and Tversky and Kahneman's (1992) CPT. Choices are assumed to be generated by a model of "stochastic choice with deterministic preferences" (Hey and Carbone 1995), and the estimations are conducted using the structural maximum likelihood estimation procedure outlined by Harrison and Rutström (2008) (and used extensively in Chapter 1).

I also estimate using maximum likelihood the standard decision models in conjunction with the Luce and Fechner error specifications, following the empirical strategy of Dave et al. (2010). These error specifications – which are analyzed in Chapter 1 as well – include a free noise parameter that captures deviations in observed choices from the predictions of the decision model considered. The noise parameter estimates indicate the extent of stochastic errors – or the magnitude of noise – in experimental decision-making, and shed light on how well the standard decision models describe observed choices (Dave et al. 2010). These error models are also estimated using the structural maximum likelihood estimation technique.

While experimental studies generally only consider one or two decision models and one of the above-mentioned error stories (for example, Harrison et al. 2007, Dave et al. 2010), I conduct estimations involving a total of six decision models – EUT-CRRA, EUT-EP, PU, RDU-CRRA, RDU-EP and CPT[5] – in combination with both the Luce and Fechner error models. The estimation of the preference and noise parameters in such a wide range of models enables a more comprehensive and complete characterization of risk attitudes, and considering a range of decision models is essential for the determination of true risk preferences.[6,7]

In the analysis so far, the noise parameter has been assumed to be constant across the entire sample of participants. This is a consequence of the traditional representative agent assumption that each subject in the experiment has the same preferences as well as the same noise parameter – each participant's choices under risk are generated by the same latent decision-making process (that is, the same decision model with the same preference and noise parameter values). However, Dave et al. (2010) note that it is important to account for the possible heterogeneity in the noise parameter across participants. Thus, following the strategy of Dave et al. (2010), I allow the noise parameter to be a linear function of various individual characteristics of the subject, in order to account for some of the heterogeneity in the noise parameter and analyze the relationship between these characteristics and stochastic errors in decision-making. These characteristics include age, gender, education and wealth, and are the same four basic characteristics used by Dave et al. (2010) in a similar analysis. Therefore, I conduct estimations once again involving the six decision models mentioned above in combination with both the error specifications, but

[5] These refer to expected utility theory constant relative risk aversion, expected utility theory expo-power, two-part power utility, rank-dependent utility constant relative risk aversion, rank-dependent utility expo-power and cumulative prospect theory, respectively.

[6] For example, if the decision-making of participants in experiments that are analyzed using only EUT models (as if often the case) is in fact better described by non-EUT models, this could yield incorrect estimates of the degree of risk aversion which do not reflect true risk attitudes (Humphrey and Verschoor 2004a).

[7] Hereafter, I refer to all the parameters in a decision model other than the noise parameter as preference parameters, and refer to the noise parameter independently.

in which the noise parameter varies by individual characteristics (while the risk preference parameters are constant across subjects). It is important to note that while there are a number of experimental studies which analyze the correlates of preference parameters in decision models, as done in Chapter 2, there are relatively few studies which account for heterogeneity in the noise parameter.[8] The maximum likelihood strategy used for these estimations is similar to that described in Section 2.3 of Chapter 2, and builds on the structural maximum likelihood estimation procedure outlined by Harrison and Rutström (2008).

In addition to estimating the magnitude of the noise parameter in Luce and Fechner error models, I provide graphical representations of the extent of noise in decision-making implied by these noise parameter estimates. If the noise parameter estimates imply substantial randomness in choice (that is, significant deviations from the predictions of the standard models), it could indicate the presence of intransitive choice patterns due to multiple switching or reverse switching (from the riskier to the safer choice) in the series of questions in each MPL decision problem. Such intransitive choice patterns would be inconsistent with all of the standard decision models considered, and indeed with any well-defined preferences over risk. I investigate the extent, and cause, of such inconsistent patterns in the data. In particular, I analyze three possible causes of the observed inconsistencies in choice (as noted by Barr 2007) – indifference between alternatives, random errors in decision-making and lack of comprehension of the decision problems.

To determine whether these patterns are caused by indifference, I analyze the

[8] In this analysis, I include explanatory variables for the noise parameter but assume the preference parameters to be constant, because the focus of this analysis is to test the link between individual characteristics and errors in experimental decision-making, and not to investigate the determinants of risk aversion or the phenomenon of asset integration (which are analyzed in Chapter 2). Further, doing so would also reduce the number of parameters to be estimated, enabling the circumvention of convergence problems in the maximum likelihood estimation procedure and the estimation of a greater number of decision and error models. Additionally, as explained in detail in later sections, if the risk preference parameters are not assumed to be constant across individuals, it would be incorrect to compare the extent of noise in decision-making (across individuals) through a direct comparison of noise parameter estimates. In Appendix B of this chapter, I propose a normalization procedure which enables the direct comparison of the noise parameter across contexts in which preference parameters differ.

extent of the range of questions over which the multiple switching occurred. Additionally, to explore the role of poor understanding, I follow the strategy of Jacobson and Petrie (2007) and use Probit and Ordered Probit regressions to evaluate whether the presence and extent of inconsistent choices vary with individual characteristics (namely, age, gender, education and wealth). Evidence of a negative relationship between inconsistencies and education (which is assumed to proxy for cognitive ability), could indicate that participants did not appropriately understand the MPL decision problems.

However, inconsistencies in choice could also arise if participants understand the decision problems and their true decision-making process is transitive, but they make random ("trembling hand") errors in decision-making. Therefore to evaluate whether the intransitive choice patterns are a result of random errors that cause deviations in observed choices from a true transitive decision-making process, or due to a true decision-making process that is intransitive (which would be indicative or poor comprehension), I use the constant error rate analysis developed by Harless and Camerer (1994).

Further, I estimate – using maximum likelihood – two distinct alternatives to the standard decision models. These two alternative models differ significantly from the standard models as, unlike the standard models, they generate intransitive choice patterns that are not in line with any well-defined preferences over risk. In order to test whether these two alternative decision models fit the data better than the standard models, and thus better represent the latent decision-making process of participants, I utilize Vuong tests and mixture model specifications (as described by Harrison and Rutström 2009), which are commonly used to compare non-nested models in the context of experimental decision-making.

3 DATA

The dataset used in this study contains the choices of participants in an experiment conducted in Brazil, in addition to socioeconomic and demographic information on the participants collected in conjunction within the experiment. The experiment design is described in Section 3.1, while the data on participant characteristics is described in Section 3.2.

3.1 Experiment design

The experiment involved 266 participants from the Ceará state in Brazil. It was conducted in August 2008 over ten sessions in two cities of this state – the capital Fortaleza and the city of Limoeiro do Norte within the Jaguaribe River Valley. The experiment was designed and conducted by researchers from Columbia University's Center for Research on Environmental Decisions, as part of a larger field experiment focusing on water management and climate variability in the region.[9,10]

Ceará is located in Brazil's relatively underdeveloped semi-arid Northeast and contains a large rural population, which is involved mostly in agricultural production (Broad et al. 2007). Recurring drought has led to low agricultural productivity and has damaged the state's economy, causing persistent poverty and vulnerability, particularly among the rural population; the rural population is poor even by local standards, with an average income less than half that implied by the national minimum wage (Broad et al. 2007). Thus, the experiment was conducted in a region where agricultural production is central to the economy and the income of most inhabitants is linked – directly or indirectly – to agricultural activities, but where rainfall and agricultural output is uncertain (Ferreira and

[9] I am very grateful to Daniel Osgood, Maria Velez and Alexander Pfaff for sharing this experimental dataset.

[10] Pfaff and Vélez (2011) and Pfaff et al. (2012) use data from the related water management experimental games to study the role of trust in the water reallocation process, and to analyze how unequal information affects the bargaining for resource allocation in this region.

Lanjouw 2001). As noted by Jacobson and Petrie (2009), such a setting is appropriate to study decisions relating to risk and insurance purchase, since individuals in this region routinely face risky choices, and risk management and insurance purchase decisions have crucial implications for individual well-being.

Each of the ten sessions of the experiment involved 20-30 participants, and sessions were conducted in universities, schools and different public institutions. The sessions conducted in the capital city of Fortaleza involved mostly college students from local families, in addition to university staff and officers from public institutions; on the other hand, the subjects in Limoeiro do Norte – within the more rural and agricultural Jaguaribe River Valley region – were mainly farmers. Recruitment of subjects was conducted through local contacts who advertised for any person older than 18 years.

Like the Ethiopian experiment considered in Chapters 1 and 2, the experiment analyzed in this chapter was designed to understand the decision-making process involving the purchase of different insurance products, as well as to evaluate the risk preferences of individuals. It included seven insurance decision problems (framed in the gain domain), each of which involved eliciting responses regarding the willingness to pay for insurance products in different risky scenarios. In eight of the ten sessions, involving a total of 212 subjects, participants responded to only one of the seven decision problems, while in the remaining two sessions, 54 participants (in total) responded to three decision problems each. Each decision problem used the Multiple Price List elicitation technique, in which participants choose their preferred option from a pair of lotteries for a series of such pairs (questions). In the context of the experiment analyzed in this chapter (hereafter, the Brazilian experiment), participants were presented with a series of 20 increasing prices in a table (one per row), and had to indicate whether they would prefer to buy an offered insurance contract ("Yes") or not ("No") at each price. Thus, in the Brazilian experiment, the MPL procedure elicits subjects' willingness to pay for insurance contracts by eliciting these binary responses for a series of 20 ordered prices in each decision problem.[11] Versions of the MPL procedure have been extensively used to elicit willingness to pay for commodities (for example, Kahneman et al. 1990), risk attitudes (for example, Holt and Laury 2002, Harrison et al. 2005) and discount rates (for example, Coller and Williams 1999). Andersen et al. (2006b) note that the Multiple Price List experimental procedure is relatively easy to implement, and it is also clear to participants that truthful responses in MPL problems are in their best interests.

The experimental setup was as follows. For each decision problem, participants were presented two boxes, labelled Box A and Box B. Each box contained a particular number of red and blue balls. Box A always contained more red balls than Box B, as well as more red balls than blue balls; on the other hand, Box B always contained more blue balls than Box A, and more blue balls than red balls. First, the enumerator blindly selected a ballot from a bag which contained a number of ballots, some with the letter A and some with the letter B – this determined whether Box A or Box B would be used. Then, the enumerator blindly drew a ball from the box (A or B) indicated by the selected ballot. The payment to participants depended, in part, on the colour of the ball blindly selected by the enumerator. Regardless of which box the ball was selected from, if a red ball was selected the participant would receive 8 Brazilian real (R$8), while if a blue ball was selected the participant would get R$16.[12] The picks of the ballot from the bag and the ball from the relevant box were common for all participants in a session.

The payment to subjects also depended on their decisions concerning whether to buy a contract whose payout was based on the box selected. The contractual

[11] Similarly, the experiment conducted by Barr and Packard (2005) in Peru also utilizes the MPL procedure to elicit choices on whether or not to purchase insurance at a series of different prices.

[12] In August 2008, the exchange rate for the US dollar to Brazilian real was approximately 1.60; thus R$8 and R$16 were equivalent to US$5 and US$10, respectively. In addition, all participants were paid a show-up fee of R$10. The national daily minimum wage in 2008 was approximately US$8.5 – therefore, the potential earnings in the experimental decision problems represented a significant amount of money, particularly for the rural population of the region, whose average income is less than half that indicated by the national minimum wage (Broad et al. 2007). As a result, the monetary prizes in the experiment are expected to be high enough to provide sufficient financial incentive for participants to answer truthfully in the decision problems (Falco 2012).

description was as follows. If the subject chose to buy a contract at a given price, he paid the price for the contract indicated in that question but would then also receive an additional payment if Box A was selected (recall that Box A always had more red balls than Box B) – specifically, the subject would receive an additional R$8, the difference between the blue ball (good outcome) and the red ball (bad outcome) payout, if Box A was selected. Enumerators stressed on the fact that this additional payment depended only on the box chosen, and not on the colour of the ball picked from the box; that is, participants would always always get paid for the ball (R$16 for blue, R$8 for red), regardless of whether they purchased the contract or not. Further, it was made clear that if a subject chose to buy the contract at a given price and Box B was selected, then he would still pay the price for the contract indicated in the question and receive the payment based on the ball colour, but would get no additional payment (from the contract). While this offered contract was essentially an insurance contract, it was only referred to as a "contract" and not an "insurance contract" when the experimental setup was described to participants.[13] For all seven decision problems, participants were presented a table of the form shown in Table 1, and had to decide whether to buy the above-described contract at each of the 20 prices.

The lowest price at which the contract was offered was R$0.40 (in question 1), and the price increased in increments of R$0.40 from question 1 to question 20. Thus, the highest price at which the contract was offered was R$8 (in question 20), which was the same amount as the additional insurance payment provided by the contract (if Box A was chosen). As the question numbers increased from 1 to 20 in a decision problem, the expected value of the "No" insurance option remained the same, while the expected value of the "Yes" insurance option decreased (as the price of insurance increased); on the other hand, the variance of both the "Yes" and "No" insurance prospects remained the same in all questions of a decision problem, with the variance of the "No" insurance option

[13] However, in the analysis conducted in this chapter, I use the terms "contract" and "insurance contract" interchangeably.

Table 1. *Table presented to participants in experimental decision problems*

Contract Price (R$)	Buy Contract?
0.40	Yes or No
0.80	Yes or No
1.20	Yes or No
1.60	Yes or No
2.00	Yes or No
2.40	Yes or No
2.80	Yes or No
3.20	Yes or No
3.60	Yes or No
4.00	Yes or No
4.40	Yes or No
4.80	Yes or No
5.20	Yes or No
5.60	Yes or No
6.00	Yes or No
6.40	Yes or No
6.80	Yes or No
7.20	Yes or No
7.60	Yes or No
8.00	Yes or No

always greater than that of the "Yes" insurance prospect. Additionally, for all 20 questions, subjects were not given the option of indicating indifference between purchasing and not purchasing the insurance contract.[14] The experiment setup – as well as the contract design – for each of the seven decision problems was identical and as described above; however, the number of A and B ballots in the bag, or the number of red and blue balls in the two boxes, or both, varied between the decision problems. The details of these aspects of the seven decision problems are provided below.

Decision Problem A: In this decision problem, there were two ballots in the bag, one with the letter A and the other with the letter B. Further, each box had five balls – in Box A all five balls were red and in Box B all five balls were blue. Participants were also informed that this implied that there was a "50%, or one out of two, chance the ball would come from Box A and a 50%, or one out of

[14] Andersen et al. (2006b) note that very few existing implementations of the Multiple Price List experimental design permit subjects to indicate indifference between alternatives.

two, chance the ball would come from Box B."

Decision Problem $B1$: In this decision problem, there were four ballots in the bag, three with the letter A and one with the letter B. Further, each box had five balls – in Box A all five balls were red and in Box B all five balls were blue. Participants were also informed that this implied that there was a "75%, or three out of four, chance the ball would come from Box A and a 25%, or one out of four, chance the ball would come from Box B."

Decision Problem $B2$: In this decision problem, there were four ballots in the bag, one with the letter A and three with the letter B. Further, each box had five balls – in Box A all five balls were red and in Box B all five balls were blue. Participants were also informed that this implied that there was a "25%, or one out of four, chance the ball would come from Box A and a 75%, or three out of four, chance the ball would come from Box B."

Decision Problem $C1$: In this decision problem, there were two ballots in the bag, one with the letter A and the other with the letter B. Further, each box had five balls – in Box A, four balls were red and one ball was blue, while in Box B, four balls were blue and one ball was red. Participants were also informed that this implied that there was a "50%, or one out of two, chance the ball would come from Box A and a 50%, or one out of two, chance the ball would come from Box B."

Decision Problem $C2$: In this decision problem, there were two ballots in the bag, one with the letter A and the other with the letter B. Further, each box had five balls – in Box A, three balls were red and two balls were blue, while in Box B, three balls were blue and two balls were red. Participants were also informed that this implied that there was a "50%, or one out of two, chance the ball would come from Box A and a 50%, or one out of two, chance the ball would come from Box B."

Decision Problem $D1$: In this decision problem, there were two ballots in the bag, one with the letter A and the other with the letter B. Further, Box A had three balls and Box B had seven balls – in Box A, all three balls were red, while in Box B, five balls were blue and two balls were red. Participants were also informed that this implied that there was a "50%, or one out of two, chance the ball would come from Box A and a 50%, or one out of two, chance the ball would come from Box B."

Decision Problem $D2$: In this decision problem, there were two ballots in the bag, one with the letter A and the other with the letter B. Further, Box A had seven balls and Box B had three balls – in Box A, five balls were red and two balls were blue, while in Box B, all three balls were blue. Participants were also informed that this implied that there was a "50%, or one out of two, chance the ball would come from Box A and a 50%, or one out of two, chance the ball would come from Box B."

Thus, like the insurance decision problems of the Ethiopian experiment, these decision problems also involved insurance purchase decisions and featured a similar two-stage probabilistic structure – first, the box to be used was randomly selected by drawing a marked ballot from a bag, and then a ball was blindly picked from the relevant box. However, unlike the insurance problems of the Ethiopian experiment, all seven decision problems in this experiment were presented in the gain, rather than loss, framework. In the context of index insurance, the box to be used could be thought of as the index and the ball chosen is comparable to the individual outcome. Further, as in an index insurance contract, the contractual payouts in the decision problems only depend on the box (A or B) chosen (that is, the index), and not on the ball picked (that is, the individual outcome). Additionally, in a real-world agricultural context, the box could be thought of as the weather, with Box A representing poor weather and Box B representing good weather. The ball could be likened to the actual yield on an individual plot, with a blue ball representing a good year (in terms of yield) for the owner and a red ball representing a bad year. Bad (good) weather is likely to lead to a bad (good) year for an agricultural producer, but this is not always the case (since in some of the decision problems there were red balls in Box B and blue balls in Box A).

Thus, the contract offered in the experimental decision problems enabled participants to purchase insurance against the bad index (Box A being chosen), but not necessarily against the bad individual outcome (red ball being picked). It is important to note, however, that unlike in the Ethiopian experiment, participants were not provided with this contextual, real-world description of the experimental setup and the insurance contract. That is, unlike the insurance decision problems of the Ethiopian experiment, the decision problems of the Brazilian experiment were framed in the abstract – thus, in the terminology proposed by Harrison and List (2004), these decision problems comprise an *artefactual field experiment*, a conventional lab experiment framed in the abstract, but conducted with a non-standard subject pool of developing country inhabitants.

In decision problems A, $B1$ and $B2$, all the balls in Box A were red and all the balls in Box B were blue. Thus, if Box A was chosen, it was certain that a red ball would be selected and participants would receive the bad outcome payout of R$8 (not accounting for the net payment from the insurance contract); additionally, if Box B was chosen, it was certain that a blue ball would be selected and subjects would receive the good outcome payout of R$16. As a result, in these decision problems, the insurance contract – which offered insurance against Box A being chosen – effectively provided insurance (of R$8) against a red ball being selected. In other words, since a red ball was selected whenever Box A was chosen, the contract provided insurance against the bad individual outcome (a red ball being selected). Thus, it was effectively a traditional indemnity insurance contract that provided full cover (by paying the entire difference of R$8 between the good and bad outcomes), despite the two-stage probability structure of the decision problems. In these decision problems, the income from the insurance contract accurately reflected the incurred loss, and there was no basis risk.

On the other hand, in decision problems $C1$ and $C2$, there were some blue balls in Box A and some red balls in Box B. Thus, there were four possible outcome states (each consisting of a box and a ball) – Box A, red ball; Box A, blue ball; Box B, red ball; Box B, blue ball.[15] Unlike in decision problems A, $B1$ and $B2$, the contract offered was an index insurance contract involving basis risk – that is, the risk that the income from insurance will not accurately reflect the incurred loss – as the income from the contract was dependent on the index (box) and did not always match the individual outcome (the ball colour). That is, it was possible that the bad index (Box A) was chosen but the good outcome (blue ball) was picked – this represented the upside basis risk whereby, if the subject chose to buy the contract, he would receive the blue ball payment (R$16) in addition to the insurance payment of R$8 (since Box A was selected). Similarly, if Box B was chosen but the red ball was picked, then the subject received the bad individual outcome payment of R$8 (since the red ball was selected), but received no payment from the insurance contract. This represented the downside basis risk – even if the subject paid to purchase the index insurance contract at a particular price, he was not guaranteed to receive the insurance payment in the case of a bad individual outcome. Thus, decision problems $C1$ and $C2$ involved both upside and downside basis risk.

In decision problem $D1$, Box A had all red balls but Box B had some red balls as well. Therefore, there were three possible outcome states – Box A, red ball; Box B, red ball; Box B, blue ball – and the contract offered was an index insurance contract involving downside basis risk, but no upside basis risk. It was possible that there was no additional income from the insurance contract even when the individual outcome was bad (that is, when Box B was chosen but a red ball was picked). Conversely, in decision problem $D2$, Box B had all blue balls but Box A had both red and blue balls. Thus, there were three possible outcome states – Box A, red ball; Box A, blue ball; Box B, blue ball – and the index insurance contract in this decision problem involved upside, but not downside, basis risk.

[15] In decision problems A, $B1$ and $B2$, there were only two possible outcome states – Box A, red ball; Box B, blue ball.

Decision problems A, $C1$, $C2$ and $D1$ were conducted in separate sessions as the second decision problem of another series of problems run by the researchers, which explored the effect of contracts on the allocation of resources.[16] 46, 58, 42 and 66 participants were involved in decision problems A, $C1$, $C2$ and $D1$, respectively. On the other hand, decision problems $B1$, $D2$ and $B2$ were conducted – in that order – in two sessions featuring all three (and no other) decision problems. Thus, the 54 participants (in total) who were involved in these two sessions each responded to all three decision problems. This implies that the dataset contains 374 responses to decision problems (from 266 individuals), where each response involved answers to 20 binary questions. Therefore, the experimental dataset includes a total of 7,480 (374×20) binary choices.

After participants had made their decisions ("Yes" or "No") about buying the contract at each of the 20 prices listed, one of the decision problems in the session was randomly selected, and one of the 20 questions (contract prices) within the chosen decision problem was randomly selected to be played out for real money, with participants being paid according to the colour of the ball picked and whether they chose to buy the contract at that price. The decision problem and the particular question to be played out for real money were randomly selected by enumerators who blindly picked from numbered envelopes in front of all the participants in the session. Harrison and Rutström (2008) note that this *random lottery payment procedure*, in which one lottery is randomly chosen to be played out for real money, is commonly used in such Multiple Price List experiments, and motivates participants to consider each choice carefully as if it were for real money (while at the same time making the experiment relatively economical to conduct); further, Charness et al. (2013) note that it provides sufficient incentive for participants to reveal preferences truthfully and is the only inventive-compatible way to implement the MPL procedure (without evoking wealth effects).

Further, to aid participant understanding, a practice round was conducted before

[16] Each of these four decision problems was conducted in two sessions.

participants made their choices in the decision problems; in addition, choices and payoffs were described orally with the help of visual aids, and randomization devices were physical, transparent and generated salient probabilities.[17] Harrison and Rutström (2008) note that such MPL problems represent a relatively transparent method to elicit risk preferences, with clear incentives for participants to answer truthfully.

3.2 Data on participant characteristics

At the end of each experimental session, participants also filled out a survey questionnaire. Thus, for each participant in the experiment, there exists data on certain important demographic and socioeconomic characteristics. Data on these characteristics is used, in Section 4.2, to analyze the correlates of errors and inconsistencies in experimental decision-making. In particular, I utilize the survey data on age, gender, education and asset ownership; further information on these variables is provided below.

In order to capture the educational attainment of subjects in the experiment, I use a binary dummy which equals one if the subject was currently enrolled in college – or was a college graduate – at the time of the experiment, and zero otherwise.[18] Gender is also represented by a binary dummy, which equals one if the subject was female. The survey dataset also contains data on the ownership of various assets – specifically, it includes binary variables providing information on the ownership of land, a house, an automobile (car or motorcycle), a washing machine and a personal computer.[19] However, a binary variable capturing ownership of a single asset may not appropriately discriminate between wealthier and poorer participants. Additionally, the households of different subjects may hold wealth in different forms. Therefore, to appropriately capture subject wealth,

[17] The use of physical and transparent randomization devices, rather than computers, helps mitigate the concern that subjects might not believe that the random process is actually fair (Harrison et al. 2010).

[18] This dataset does not contain direct measures of quantitative literacy and understanding of the decision problems, unlike the dataset associated with the Ethiopian experiment.

[19] Each binary variable equals one if the subject owned that particular asset and zero otherwise.

Table 2. *Weights assigned to each asset indicator in the PCA wealth index*

Asset	Weight in PCA Wealth Index
Land	0.327
House	0.336
Automobile	0.506
Washing machine	0.530
Personal computer	0.494

I create a composite wealth measure, using the five binary asset ownership variables mentioned above. This wealth measure (or index) is constructed to be the first principal component arising from a principal components analysis (PCA) of these five binary asset ownership variables.[20]

Table 2 presents the weights from the principal eigenvector, which are used to construct the PCA wealth index. As might be expected, all weights have the same sign, indicating that the five asset ownership variables are jointly positively correlated.

Using the asset ownership variables and corresponding weights provided by the PCA, the first principal component (that is, the PCA wealth index) for individual a is specified as:

$$\begin{aligned} PCAwealth_a = 0.327 \times Land_a + 0.336 \times House_a + 0.506 \times Automobile_a \\ + 0.530 \times WashingMachine_a + 0.494 \times PersonalComputer_a \end{aligned} \tag{1}$$

The first principal component provides the linear combination (of the variables) with maximum variance, thus enabling better discrimination between wealthier and poorer subjects; further, by accounting for the ownership of various different assets – rather than a single asset – it provides a multi-dimensional and more complete measure of overall wealth (Filmer and Pritchett 2001). Filmer and

[20] Moser and Felton (2007) and Howe et al. (2008) provide a detailed analysis of PCA as a tool for constructing wealth indices, including some of the advantages and disadvantages. PCA is not scale invariant and so the analysis is conducted using the correlation matrix. This is equivalent to using standardized variables.

Table 3. *Summary statistics*

Variable	No. of Observations	Mean	Standard Deviation
Age (years)	372	24.03	6.144
Gender (1 if female, 0 if male)	374	0.476	0.500
Education (1 if college educated, 0 if not)	374	0.709	0.455
PCA wealth index	374	1.047	0.628
Land	374	0.179	0.384
House	374	0.743	0.437
Automobile	374	0.532	0.500
Washing Machine	374	0.404	0.491
Personal Computer	374	0.519	0.500

Pritchett (2001), using survey data from rural India to construct and analyze a similar wealth index, find that a PCA wealth index generated from the first principal component using discreet asset ownership indicators provides a reliable proxy for the long-run economic status of households. Additionally, Vyas and Kumaranayake (2006) note that wealth measures based on asset data – rather than income or consumption data – involve lower measurement error. Such PCA indices have been used extensively in published studies as measures of wealth in the context of developing countries (for example, Filmer and Pritchett 2001, Schellenberg et al. 2003), including in analyses involving data from Brazil (for example, Barros and Victora 2005, Vyas and Kumaranayake 2006).

The four characteristics – age, gender, college education and PCA wealth – are the explanatory variables used to describe errors and inconsistencies in experimental decision-making. The summary statistics for these variables are provided in Table 3.[21]

Table 3 indicates that the average age of participants in the experiment is 24 years – this is considerably lower than the average age of subjects involved in the Ethiopian experiment (45 years). This is explained largely by the fact that a majority of participants (around 60%) in the Brazilian experiment were enrolled college students. Further, the summary statistics indicate that nearly 71% of

[21] Note that the college education binary variable is highly correlated with the occupation of participants, since 214 (or 80%) of the 266 subjects are either college students or farmers.

participants were either currently enrolled in, or had completed, college at the time of the experiment. According to the formal education system in Brazil, this implies that 71% of experiment participants had completed high school and thus received at least 11 years of formal education (Pfaff et al. 2012). In comparison, participants in the Ethiopian experiment had received only four years of formal schooling, on average. Additionally, around 48% of participants in the Brazilian experiment were female, as compared to 30% in the Ethiopian experiment.

The PCA wealth index has a sample mean of around one. Almost three-quarters of the participants lived in a house that they owned (obtained either through buying, building or inheriting), while only 18% owned land. Further, automobile and personal computer ownership were both around 50%, while 40% of subjects owned a washing machine.

4 EMPIRICAL ANALYSIS

4.1 Estimation of standard decision models

In this section, I build on the analysis conducted in Chapter 1 and estimate the various decision models detailed in that chapter, using the experimental data from Brazil. As the starting point of this empirical analysis, I estimate an EUT specification with a CRRA utility function. The CRRA utility function is given by[22]:

$$U(x) = \begin{cases} \frac{x^{1-r}}{1-r} & \text{if } r \neq 1 \\ \ln(x) & \text{if } r = 1 \end{cases} \tag{2}$$

where x is the lottery prize and r is the coefficient of relative risk aversion: $r < 0$ corresponds to risk loving behaviour, $r = 0$ to risk neutral and $r > 0$

[22] The notation used in this chapter in relation to decision models and preference parameters is the same as that used in Chapter 1, unless otherwise specified.

to risk averse.[23,24] Additionally, under EUT, the decision-maker weights each possible outcome $k_i \in \{1, \ldots, K\}$ in lottery i using the objective probability p_{k_i} associated with the outcome, and so expected utility from lottery i (EU_i) is the sum of the probability weighted utility of each outcome in the lottery:

$$EU_i = \sum_{k_i=1}^{K} p_{k_i} U_{k_i} \tag{3}$$

For each question q in each decision problem d, a participant either chose to buy insurance ("Yes") or not to buy insurance ("No"). Thus, in other words, the participant chose between an option that involved less risk (insurance purchase) – that is, less variation in potential outcome prizes – and one which involved more risk (no insurance purchase), or greater variation in potential outcome prizes. Thus, for each question, $i \in \{Y, N\}$, where Y denotes the option involving insurance purchase and N denotes the prospect involving no insurance purchase.

To fit this model to the data and estimate the parameter r, I use a structural model combined with a maximum likelihood estimation technique, following the estimation strategy pioneered by Camerer and Ho (1994) and detailed in the work of Harrison and Rutström (2008). The structural estimation methodology – applied to the MPL decision problems of the Brazilian experiment – is as follows. First, the expected utility $EU_i^{d,q}$ from each potential lottery choice $i \in \{Y, N\}$ in each question $q \in \{1, \ldots, 20\}$ of each decision problem $d \in \{A, B1, B2, C1, C2, D1, D2\}$ is calculated according to Equation (3), assuming that utility is defined by the CRRA function in Equation (2). The latent index

23 The Arrow-Pratt measure of relative risk aversion, or coefficient of relative risk aversion, is defined as $\frac{-xU''(x)}{U'(x)}$.

24 In all the estimations in this chapter, the show-up fee (of R$10) in the experiment is excluded from the lottery prize x (as done in Chapter 1); this is common practice in empirical analyses of experimental data, and assumes that participants do not integrate their show-up fee with earnings from the experimental decision problems (Harrison and Rutström 2008). This also implies that in reference-dependent models (such as those involving loss aversion), the reference point does not include the show-up fee. However, the results remain substantively the same when the show-up fee is included in the measure of the lottery prize x.

$\nabla EU_{1,i}^{d,q}$ is then calculated as follows:

$$eu_i^{d,q} = \exp(EU_i^{d,q}) \tag{4}$$

$$\nabla EU_{1,i}^{d,q} = \frac{eu_i^{d,q}}{eu_Y^{d,q} + eu_N^{d,q}} \tag{5}$$

This index $\nabla EU_{1,i}^{d,q}$ (hereafter referred to as Latent Index 1), which links latent preferences to choice probabilities (and thus observed choices), is simply a two-option version of that used in Chapters 1 and 2, when analyzing data from Ordered Lottery Selection problems.[25] It has also been used extensively in analyses of MPL decision problems (for example, Holt and Laury 2002, Heinemann 2008, Dave et al. 2010). The latent index $\nabla EU_{1,i}^{d,q}$, based on latent preferences, is in the form of a probability between 0 and 1, and thus can be directly linked to the observed choices. Following Harrison and Rutström (2008), $\nabla EU_{1,i}^{d,q}$ is interpreted as the probability of a subject choosing lottery choice i in question q of decision problem d.

However, when considering choices between pairs of prospects, it is possible to use another type of latent index. This index, outlined by Harrison and Rutström (2008), is given by:

$$\nabla EU_2^{d,q} = EU_Y^{d,q} - EU_N^{d,q} \tag{6}$$

where $EU_Y^{d,q}$ and $EU_N^{d,q}$ are the expected utility values – evaluated using Equations (2) and (3) – associated with the "Yes" insurance and "No" insurance options, respectively. This index $\nabla EU_2^{d,q}$ (hereafter referred to as Latent Index 2), however, is not in the form of a probability between 0 and 1, and thus a functional transformation is required to link the index to choice probabilities (and thus observed choices). As noted by Harrison and Rutström (2008), one popular option for this statistical link function is the standard normal cumulative

[25] A detailed derivation of this structural model, including the specific assumptions underlying it, is provided in Appendix A of Chapter 1.

distribution function.[26] Using this latent index in combination with the standard normal cumulative distribution function (Φ) gives:

$$\text{Prob(choose lottery } Y) = \Phi(\nabla EU_2^{d,q}) = \Phi(EU_Y^{d,q} - EU_N^{d,q}) \qquad (7)$$

Using this "probit" link function, the probability of a subject choosing option Y is the standard normal cumulative probability up to $\nabla EU_2^{d,q} = EU_Y^{d,q} - EU_N^{d,q}$.[27] Thus,

$$\text{Prob(choose lottery } N) = 1 - \Phi(\nabla EU_2^{d,q}) = \Phi(-\nabla EU_2^{d,q}) = \Phi(EU_N^{d,q} - EU_Y^{d,q}) \qquad (8)$$

Therefore, for both Latent Index 1 and 2, the index can be related to the probabilities of observing Y and N choices in a particular question. When using Latent Index 1, $\nabla EU_{1,y_a^{d,q}}^{d,q}$ gives the probability of observing choice $y_a^{d,q}$ (which is either Y or N) for individual a in question q of decision problem d. Therefore, using Latent Index 1, the log-likelihood of the observed responses of all participants in all 20 questions of all seven decision problems – conditional on the EUT-CRRA specification representing the true decision-making process of subjects – is:

$$\ell^{EUT} = \ln L^{EUT}(r; \boldsymbol{y}) = \sum_d \sum_{a=1}^{N_d} \sum_{q=1}^{20} \ln(\nabla EU_{1,y_a^{d,q}}^{d,q}) \qquad (9)$$

where N_d is the number of participants involved in decision problem d and $\boldsymbol{y}$ is the vector of observed choices in all questions of all decision problems. Using

[26] Another popular option for the link function is the logistic cumulative distribution function, as used by Harrison and Rutström (2009). However, as noted by Train (2009), Latent Index 2 in combination with the logistic link function collapses to Latent Index 1 in the two-option case.

[27] Harrison et al. (2007) and Harrison et al. (2011) also utilize the same latent index, along with the standard normal link function, in their structural maximum likelihood estimations of risk preferences.

Latent Index 2, the corresponding log-likelihood is:

$$\ell^{EUT} = \ln L^{EUT}(r; \boldsymbol{y}) = \sum_{d} \sum_{a=1}^{N_d} \sum_{q=1}^{20} [\ln(\Phi(\nabla EU_2^{d,q}) \times \mathbf{I}(y_a^{d,q} = Y)) + \ln((1 - \Phi(\nabla EU_2^{d,q})) \times \mathbf{I}(y_a^{d,q} = N))] \tag{10}$$

where $\mathbf{I}(\cdot)$ is the indicator function and $y_a^{d,q} = Y(N)$ denotes the choice of the Yes (No) insurance option by individual a in question q of decision problem d.

As indicated above, I jointly consider the choices of all individuals in all questions of all decision problems.[28] If there are no missing choices in the experimental data, this implies a total of 7480 binary choices included in these estimations. The log-likelihood function can then be maximized with respect to the core parameter r to yield the structural maximum likelihood estimate of the coefficient of relative risk aversion r.

Even though decision theories – such as EUT, RDU and CPT – are deterministic, the structural model described above does incorporate stochastic errors in decision-making[29]; however, it assumes that the extent of noise in decision-making is fixed at a certain level (see Appendix A of Chapter 1). When preference parameters are assumed to be constant across individuals, these errors account for heterogeneity in the observed choices of participants. Harrison and Rutström (2008) note that an important extension of the core model is to allow for the estimation of the magnitude (or extent) of errors in decision-making – rather than assuming a fixed magnitude of noise – as discrete choice is highly stochastic.

[28] Note that here I assume that all the binary choices in the dataset are independent, and generated by a single model of decision-making under risk (following the strategy of Harrison et al. 2007, Harrison and Rutström 2008).

[29] It is important to allow for a stochastic element in decision-making, otherwise a single observation in violation of a particular decision theory would be sufficient to reject it (Starmer 1999). Further, numerous studies have found evidence of substantial errors in experimental decision-making; in addition, the one-shot nature of the experiment as well as the unfamiliarity of the Brazilian participants with such decision problems is likely to induce some noise in revealed preferences, and thus any model used to describe experimental data (that is, the data generating process) should be specified stochastically (that is, with allowance for observed choices to deviate from true preferences) (Starmer 1999, Humphrey and Verschoor 2004a).

As done in Chapter 1, I consider two important "white noise" stochastic error specifications to estimate the magnitude of noise in experimental decision-making – the Luce (1959) error specification and the Fechner (1966) error specification. These error models (as well as the models used earlier), represent models of "stochastic choice with deterministic preferences", since the true preferences are deterministic but the observed choices are stochastic due to the presence of "white noise" error (Hey and Carbone 1995).[30] That is, the errors allow observed choices to deviate from true preferences.

These error models have been used extensively in the experimental literature on decision-making under risk (for example, Holt and Laury 2002, Dave et al. 2010, Harrison et al. 2007). They place restrictions on the violations or errors made by participants, and provide useful statistical information and instructions (Wilcox 2008). The Luce error specification is used in conjunction with Latent Index 1, and provides the following modified index:

$$\nabla EU^{d,q}_{iLuce} = \frac{\exp(EU^{d,q}_i/\mu)}{\exp(EU^{d,q}_Y/\mu) + \exp(EU^{d,q}_N/\mu)} \tag{11}$$

Similarly, the Fechner error specification is used in conjunction with Latent Index 2, and provides the following modified index:

$$\nabla EU^{d,q}_{Fech} = \left(\frac{EU^{d,q}_Y - EU^{d,q}_N}{\mu}\right) \tag{12}$$

In both error models, μ represents the error term or noise parameter, and captures the extent of randomness in choice.[31] As $\mu \rightarrow 0$, the specification collapses to the deterministic EUT model, where the choice is strictly determined by the

[30] While the structural estimation model described earlier does incorporate errors as well, I follow the classification of Harrison and Rutström (2008) and distinguish the "error" models described here from the models described earlier (which assume a fixed extent of noise in decision-making). Further, Appendix A and B of Chapter 1 provide detailed derivations of both types of models.

[31] Specifically, the μ parameter determines the variance of the error term in the model specification (provided in Appendix B of Chapter 1). However, following Harrison and Rutström (2008) and Harrison et al. (2011), I refer to this parameter as the "error" or "noise" parameter, since it determines the extent to which choice probabilities depart from true preferences and thus the magnitude of stochastic errors in experimental decision-making.

expected utility of the two lotteries, and the option providing greater utility is always chosen (Harrison and Rutström 2008). Larger values of μ imply larger errors in the decision-making process and a greater deviation of observed choices from true (deterministic) preferences, or a greater extent of deviation in choice probabilities from the expected utilities of the lotteries, assuming that the true preferences of participants are described by the decision model used to evaluate EU (in this case, EUT with a CRRA utility function).[32] Thus, the deviation in μ from the no error value of 0 provides an indication of the extent of stochastic errors, and any choices that are inconsistent with the predictions of the decision theory under consideration are attributed to these errors. These errors may arise due to carelessness or miscalculations on the part of individuals when evaluating alternatives in a question, due to imperfect access to their true utility function, or as a result of subjects misunderstanding the nature of the decision problem (Hey 1995, Loomes 2005). Further, from Equations (11) and (12), it is clear that when $\mu = 1$, $\nabla EU^{d,q}_{iLuce} = \nabla EU^{d,q}_{1,i}$ and $\nabla EU^{d,q}_{Fech} = \nabla EU^{d,q}_{2}$. Thus, the model considered previously (and the associated latent indices, given in Equations (5) and (6)) is nested within this error specification and incorporates stochastic errors in decision-making (that is, μ is not assumed to be 0), but fixes the magnitude (or extent) of these errors (at $\mu = 1$).

The modified latent indices (incorporating the error parameters), which are given in Equations (11) and (12), are used to construct log-likelihood functions as described in Equations (9) and (10). The log-likelihood functions can then be maximized with respect to the preference parameter r and the noise parameter μ to obtain the structural maximum likelihood estimates of these parameters. The results of these maximum likelihood estimations – as well as those involving the latent indices without the error parameter (that is, assuming $\mu = 1$) – are

[32] In the case of the Fechner error model, the magnitude of μ determines the extent to which choice probabilities depart from the difference in expected utilities of the right and left lotteries (that is, true preferences).

Table 4. *Structural maximum likelihood estimates of risk attitudes for EUT with CRRA utility*

Coefficient	Estimate	Standard Error	p-value	Lower 95% Confidence Interval	Upper 95% Confidence Interval	Observations
A. Latent Index 1: no error parameter						
r	0.360	0.012	0.000	0.336	0.383	7452
B. Latent Index 1: Luce error parameter						
r	-1.377	0.171	0.000	-1.712	-1.042	7452
μ	72.64	31.93	0.023	10.06	135.2	7452
C. Latent Index 2: no error parameter						
r	0.583	0.012	0.000	0.561	0.606	7452
D. Latent Index 2: Fechner error parameter						
r	-1.361	0.175	0.000	-1.704	-1.019	7452
μ	117.6	52.86	0.026	14.00	221.2	7452

The log-likelihoods corresponding to the estimations in Panels A, B, C and D are -4328.5, -4306.1, -4337.1 and -4314.6.
Wald test p-values and 95% confidence intervals reported.

presented in Table 4.[33,34]

The estimates of the coefficient of relative risk aversion r are 0.36 and 0.58 in the EUT-CRRA specifications without the error parameter using Latent Index 1 and 2, respectively (reported in Panels A and C of Table 4).[35] Both these estimates are statistically significantly different from 0 at the 1% level, as indicated by Wald tests.[36] The structural estimation methodology used to estimate model parameters makes the traditional representative agent assumption that each subject in the experiment has the same preferences – that is, each participant's choices under

[33] I refer to the specifications that assume $\mu = 1$ as those without the error parameter, because there is no explicit free noise parameter to be estimated in these specifications (even though they incorporate stochastic errors).
[34] The maximum likelihood estimations reported in this chapter were carried out in Stata – using the Newton-Raphson and Berndt-Hall-Hall-Hausman optimization algorithms – and MATLAB (using the sequential quadratic programming optimization algorithm).
[35] Note that the number of observations included in each of the estimations is 7452, since there were a total of 28 missing responses (to binary questions) in the experimental dataset.
[36] I also conduct likelihood ratio tests in addition to Wald tests to test the various hypotheses in this chapter – in all cases the conclusions from both tests are identical.

risk are assumed to be generated by the same decision model with the same preference parameter values (see Appendix A of Chapter 1); this assumption has been incorporated in similar estimations by Andersen et al. (2006a), Harrison and Rutström (2009) and Harrison et al. (2011). Thus, the parameter estimates of the decision models can be interpreted as the parameter values for the average subject. The estimates in Table 4 indicate that, if EUT with CRRA utility is assumed to be the decision-making process generating participants' choices in the decision problems, the participants in the experiment exhibit moderate risk aversion, on average. Further, these estimates of the average r are comparable to those obtained in Chapter 1, using data from Ethiopia and a similar structural model. However, when the error parameter is included in the specification (Panels B and D), the estimated r is negative and statistically significant – this implies that when I account for the extent of stochastic errors in the decision-making process, participants exhibit risk loving behaviour in the experimental decision problems. Andersen et al. (2006a) and Charness and Viceisza (2012) also find some evidence that participants in experiments conducted in the United States and rural Senegal, respectively, exhibit risk loving behaviour. Charness and Viceisza (2012) note, however, that since risk loving behaviour has rarely been observed in experimental studies, this could indicate a lack of understanding on the part of subjects (leading to "frivolous" responses) – this possibility is further explored in Sections 4.2 and 4.3.[37]

The estimated noise parameter is relatively large in magnitude compared to the no error value of 0 – it is around 73 in the Luce error specification and 118 in the Fechner error specification. Both estimates are statistically significantly different from 0 at the 5% level, implying a rejection of the deterministic choice hypothesis ($\mu = 0$) that there are no stochastic errors in the decision-making process and that observed choices are determined solely by true preferences

[37] Additionally, it is important to note that in these specifications, the estimates of the preference parameters represent the parameter values conditional on EUT-CRRA (plus error) representing the true data-generating process. Thus, if subjects' true preferences diverge substantially from EUT-CRRA, the preference parameter estimates may not accurately reflect true risk attitudes.

(in this case, EUT with CRRA utility). Following the interpretation provided by Dave et al. (2010), the large estimated magnitude (relative to 0) of the noise parameter indicates that subjects' choices in the experimental decision problems may not conform closely to the predictions of EUT-CRRA. In other words, choice probabilities do not depend highly on the expected utilities of the lotteries, and this could be considered as evidence of a high degree of randomness in decision-making; alternatively, it could indicate that EUT-CRRA does not represent the true decision-making process of participants (and therefore reflects misspecification).[38] Further, substantial randomness in choice could be indicative of poor understanding of the complex MPL decision problems (Crosetto and Filippin 2013); indeed, Humphrey and Verschoor (2004a) note that the noise may depend on the difficulty of the decision problems conducted, with greater errors more likely in complex problems. Additionally, using Wald tests, the hypothesis of $\mu = 1$ is rejected at the 5% level for both error specifications.

The results in Table 4 highlight the importance of accounting for and estimating the magnitude of stochastic variation in the decision-making process – not accounting for the noise parameter could lead to markedly different parameter

[38] It may be a concern that, with the representative agent assumption, the noise parameter identifies heterogeneity in risk preferences across subjects rather than randomness in choice (or errors in decision-making). However, as will be considered in detail in later sections, a significant fraction of the choice patterns observed in the data are intransitive (and thus inconsistent with any well-defined preferences over risk), and these patterns are likely to be a result of poor understanding of the decision problems – this shows that there is significant randomness in choice, which in turn results in statistically significant noise parameter estimates; further, it indicates that the large noise parameter estimates (relative to the no error parameter value of 0) are not generated due to heterogeneity in risk preferences. To completely account for heterogeneity (both observed and unobserved) in risk preferences, a within-subject analysis (as conducted by Hey and Orme 1994) is required, which separately considers the choices of each participant. Analyzing the choices – and risk preferences implied by these choices – of each participant in different decision problems would enable me to conclusively determine whether subjects' choices are generated (largely) by errors or not, while appropriately accounting for heterogeneity in risk preferences. However, in the Brazilian experiment, only 54 participants were involved in more than one decision problem, and of these subjects, only 22 made consistent (transitive) choices in each of the three decision problems they responded to. Thus, there are too few subjects who were involved in multiple decision problems and made consistent choices (and too few choices per subject) to conduct a meaningful within-subject analysis. However, analyzing the choices of these 22 participants, I find that, for a particular subject, the choices in the three decision problems imply very different levels of risk aversion. Thus, given this substantial variation in implied risk preferences between the different choices of a particular subject, this within-subject analysis (which accounts for heterogeneity in risk preferences across subjects) indicates that there is significant noise in the decision-making of participants.

estimates and thus contrasting qualitative and quantitative descriptions of behaviour under risk (a point also noted by Loomes et al. 2002). That is, when μ is significantly different from 1, as is the case in the Luce and Fechner error estimations in Table 4, the structural estimations without the μ parameter – which implicitly assume that $\mu = 1$ – may yield different, and possibly incorrect, estimates of the preference parameters. This is because these estimations provide the maximum likelihood estimates of the preference parameters (in the case of EUT-CRRA, r) given that the noise parameter is fixed at 1; however, if μ is significantly different from 1, the preference parameter estimates may be very different in the specifications involving the free noise parameter. For example, the results for the EUT-CRRA specification show that the estimated r is 0.360 (indicating moderate risk aversion) when it is assumed that $\mu = 1$, but is -1.377 (indicating risk loving behaviour) when μ is considered to be a free parameter.

However, numerous studies in the experimental literature utilize structural specifications that assume $\mu = 1$, as they permit errors but do not increase the number of parameters to be estimated (for example, Harrison and Rutström 2008, Harrison et al. 2010). Therefore, throughout this thesis I conduct estimations of the various decision models using these specifications – in addition to those involving the free noise parameter – for comparative purposes, and primarily to compare the preference parameter estimates obtained in this chapter to those obtained in Chapter 1 using similar structural specifications.

Conversely, when μ is allowed to vary and is estimated jointly with the preference parameters, the maximum likelihood estimate of μ depends on the estimates of the preference parameters – different preference parameter values would yield different maximum likelihood estimates of the noise parameter. In other words, the estimate of μ varies with the estimates of the preference parameters, independently of differences in the extent of noise (errors) in decision-making. Therefore, with μ in its current form, the extent of noise in experimental decision-making cannot be compared directly through the comparison of the magnitude of μ if the preference parameter values are different – since the estimate of μ

depends also on the preference parameter values (estimates), variation in μ does not solely capture differences in noise. The comparison of the noise parameter μ across contexts in which the preference parameters differ has no ordinal or cardinal interpretation. For example, μ estimates are not comparable across experimental studies, since different samples are expected to have different risk attitudes, or across decision models that involve different preference parameters. Thus, the results of studies (for example, Dave et al. 2010, Crosetto and Filippin 2013) that utilize a comparison of noise parameter estimates to compare the extent of errors in decision-making across contexts in which the preference parameters differ could be misleading.

In Appendix B of this chapter, I propose a normalization procedure for the noise parameter (μ) which would make it comparable across these contexts. This enables the direct comparison of the extent of errors in decision-making by different groups of individuals through a comparison of the magnitude of the normalized μ. Appendix B also reports the results of the comparison of the normalized noise parameter estimates across the Brazilian and Ethiopian experiments analyzed in this thesis.

It is important to note that while the EUT-CRRA model with the Luce error parameter could not be estimated using the Binswanger-style lottery data in Chapter 1 (which included a similar number of participants and decision problems), it is estimated – with both the r and μ parameters estimated with relatively narrow confidence intervals – using the data from the Multiple Price List decision problems of the Brazilian experiment. As mentioned in Section 1, this is possibly because the MPL procedure enables a finer and more precise characterization of risk attitudes; in other words, it is possible that this procedure provides more information and greater power for the estimation of different decision models (Dave et al. 2010, Charness et al. 2013).[39] While an Ordered Lottery Selection decision problem involves a single choice between multiple options and thus only allows the classification of subjects into relatively few preference categories (in the case of the Ethiopian experiment, six), the Multiple Price List decision problems of the Brazilian experiment involve 20 binary choices and thus allow the classification of subjects into 20 preference categories.[40]

In addition to estimating the CRRA utility function, I also estimate a more flexible two-parameter expo-power (EP) utility function – developed by Saha (1993) – within the EUT framework. The EUT-EP specification, described in detail in Section 2.1 of Chapter 1, also utilizes Equation (3) to aggregate outcomes in a lottery, but uses the following utility function to evaluate each outcome:

$$U(x) = \frac{(1 - \exp(-\alpha_{EP} x^{1-r_{EP}}))}{\alpha_{EP}} \tag{13}$$

The latent indices (both with and without the error parameter) are computed and used to construct the log-likelihood functions as done for the EUT-CRRA specification. These log-likelihood functions can then be maximized with respect to the two preference parameter (r_{EP} and α_{EP}), in addition to the noise parameter in the error specifications, to obtain structural maximum likelihood estimates of these parameters. These estimates are presented in Table 5.

In the specifications without the noise parameter (Panels A and C), the estimates

[39] Further, unlike for the maximum likelihood estimations in Chapter 1 for which non-convergence of the optimization procedure was encountered, scans (as well as graphs of the log-likelihood function) do not indicate that the log-likelihood function is flat over a range of maximum likelihood parameter estimates – that is, the evidence does not indicate that multiple parameter estimates produce the same (maximized) value of the log-likelihood function, and there is sufficient information to identify unique maximum likelihood estimates. It could be that this greater information is provided by the greater precision of the MPL elicitation procedure, as compared to the Ordered Lottery Selection procedure used in the Ethiopian experiment (which is analyzed in Chapter 1).

[40] Strictly speaking, there are 21 preference categories, depending on the switch point from the Yes to No insurance option – that is, 19 categories corresponding to choice patterns with switch points in questions 2 through 20, as well as the categories corresponding to the all Yes and all No choice patterns. However, in question 20, the Yes insurance option is dominated by the No insurance option and thus a pattern with all Yes choices is a dominated one. Therefore, there are only 20 meaningful risk preference categories (a similar point is noted by Crosetto and Filippin 2013).

Table 5. *Structural maximum likelihood estimates of risk attitudes for EUT with expo-power utility*

Coefficient	Estimate	Standard Error	p-value	Lower 95% Confidence Interval	Upper 95% Confidence Interval	Observations
A. Latent Index 1: no error parameter						
r_{EP}	0.534	0.065	0.000	0.408	0.661	7452
α_{EP}	-0.389	0.156	0.013	-0.695	-0.083	7452
B. Latent Index 1: Luce error parameter						
r_{EP}	0.316	0.176	0.073	-0.030	0.662	7452
α_{EP}	-0.351	0.295	0.235	-0.930	0.229	7452
μ	5.063	1.220	0.000	2.672	7.455	7452
C. Latent Index 2: no error parameter						
r_{EP}	0.614	0.063	0.000	0.491	0.737	7452
α_{EP}	-0.415	0.186	0.025	-0.779	-0.051	7452
D. Latent Index 2: Fechner error parameter						
r_{EP}	0.305	0.165	0.065	-0.019	0.629	7452
α_{EP}	-0.327	0.257	0.202	-0.830	0.176	7452
μ	8.222	1.913	0.000	4.472	11.97	7452

The log-likelihoods corresponding to the estimations in Panels A, B, C and D are -4314.5, -4300.3, -4325.5 and -4307.9.
Wald test p-values and 95% confidence intervals reported.

of r_{EP} are around 0.55 and those of α_{EP} are around -0.4. The estimates of α_{EP} are significantly different from 0 at the 5% level in both specifications, indicating that relative risk aversion varies with the size of the lottery prize x, and thus implying rejection of the CRRA hypothesis ($\alpha_{EP} = 0$) in favour of decreasing relative risk aversion ($\alpha_{EP} < 0$). Additionally, the estimates of r_{EP} are also significantly different from 0 (at the 1% level) in both specifications, implying a rejection of the constant absolute risk aversion hypothesis ($r_{EP} = 0$) in favour of decreasing absolute risk aversion ($r_{EP} > 0$). In comparison, in Chapter 1 I obtained positive and statistically significant estimates of both EP parameters in similar specifications.

In the two error specifications (Panels B and D), the estimates of r_{EP} are once again positive and statistically significant (at the 10% level). However, the estimates of α_{EP} are no longer statistically significant, indicating that the hypothesis of CRRA ($\alpha_{EP} = 0$) cannot be rejected. The estimated Luce and Fechner error parameters are approximately 5 and 8, respectively. Both estimates are also statistically distinguishable from 0 at the 1% significance level, indicating a rejection of the hypothesis that choices are deterministic and there are no stochastic errors in the decision-making process. Instead, the large estimates (relative to the no error parameter value of 0) indicate significant deviations in observed choices from the predictions of the EUT-EP model, that is, a low dependence of choice probabilities on the expected utilities of the lotteries. Additionally, the hypothesis of $\mu = 1$ is once again rejected (at the 1% level) for both error models.

As done in Chapter 1, I deviate from the assumptions of EUT and allow for the possibility that decision-makers evaluate gains and losses in risky lotteries differently. I estimate a two-part power utility (PU) function and investigate the presence of loss aversion, whereby losses loom larger than gains in subjects' valuations of risky lotteries. This form of utility function was used by Tversky and Kahneman (1992) in their development of cumulative prospect theory, and has been used in recent empirical studies by Liu (2008) and Tanaka et al. (2010).

In this specification, the outcomes in each lottery are aggregated using objective probabilities as described in Equation (3), but the utility function used to evaluate each outcome is:

$$U(x,\chi) = \begin{cases} (x-\chi)^{\alpha_P} & \text{if } (x-\chi) \geq 0 \\ -\lambda(-(x-\chi))^{\alpha_P} & \text{if } (x-\chi) < 0 \end{cases} \tag{14}$$

Thus, utility is no longer defined over the final prize amounts (x) in the lotteries, but is defined separately over gains and losses relative to a reference point χ. As done in Chapter 1, I consider the reference point to be endogenous and estimate it (jointly with the other parameters in the model) using maximum likelihood.[41] Further, α_P is the parameter capturing the curvature of the utility function (referred to as the risk aversion parameter by Harrison and Rutström 2009) and λ is the loss aversion parameter. Latent Index 1 and 2 are used in combination with this decision model to provide the log-likelihood function, which can be maximized with respect to the three parameters (χ, α_P and λ) – in addition to the noise parameter in the error specifications – to obtain structural maximum likelihood estimates of these parameters. The results of these estimations are reported in Table 6.

In the specifications without the error parameter (Panels A and C), the reference point χ is estimated to be approximately R$4, but in the error specifications (Panels B and D) it is estimated to be around R$8. This once again highlights that estimated risk preferences could differ significantly depending on whether the magnitude of stochastic errors is accounted for or not – in other words, parameter estimates (and thus estimated risk preferences) are significantly impacted by, and are dependent on, the model specification. The lottery prizes x in the experiment range from R$0 to R$23.60; thus, as found in Chapter 1, the reference point estimates are near the lower end of the spectrum of lottery prizes. However,

[41] This is in accordance with the empirical strategy of Andersen et al. (2006a) and Harrison and Rutström (2008), and the theoretical arguments of Munro and Sugden (2003) and Falk and Knell (2004). Additionally, the advantages and disadvantages of this method are highlighted in Section 5.4 of Chapter 1.

Table 6. *Structural maximum likelihood estimates of risk attitudes for the two-part power utility function*

Coefficient	Estimate	Standard Error	p-value	Lower 95% Confidence Interval	Upper 95% Confidence Interval	Observations
A. Latent Index 1: no error parameter						
χ	4.638	0.210	0.000	4.227	5.049	7452
λ	3.605	0.309	0.000	3.000	4.210	7452
α_P	0.671	0.011	0.000	0.651	0.692	7452
$H_0 : \lambda = 1$			0.000			
B. Latent Index 1: Luce error parameter						
χ	7.826	0.333	0.000	7.173	8.480	7452
λ	4.015	0.895	0.000	2.260	5.770	7452
α_P	1.625	0.089	0.000	1.451	1.799	7452
μ	10.57	2.584	0.000	5.501	15.63	7452
$H_0 : \lambda = 1$			0.000			
C. Latent Index 2: no error parameter						
χ	4.052	0.152	0.000	3.754	4.350	7452
λ	2.199	0.198	0.000	1.812	2.587	7452
α_P	0.536	0.009	0.000	0.518	0.554	7452
$H_0 : \lambda = 1$			0.000			
D. Latent Index 2: Fechner error parameter						
χ	8.005	0.107	0.000	7.795	8.215	7452
λ	3.464	0.355	0.000	2.769	4.160	7452
α_P	1.597	0.070	0.000	1.459	1.735	7452
μ	16.04	2.699	0.000	10.75	21.33	7452
$H_0 : \lambda = 1$			0.000			

The log-likelihoods corresponding to the estimations in Panels A, B, C and D are -4260.6, -4165.9, -4279.8 and -4171.6.
Wald test p-values and 95% confidence intervals reported.

outcomes greater than R$16 only occur in the three decision problems involving upside basis risk ($C1$, $C2$ and $D2$), while prizes around R$4 and R$8 are present in every decision problem. Thus, the estimates of the reference point are around the most commonly occurring prize values. It is also important to note that in all the four estimations presented in Table 6, the estimates of χ are positive and statistically significantly different from 0 at the 1% level[42], which is expected given that all the decision problems are presented in the gain framework and there are no negative prize outcomes. The results are in line with the observation of Crosetto and Filippin (2013) that even for decision problems framed completely in the gain domain, the reference point could be endogenously determined by the outcomes in the task, and gains and losses evaluated against this reference point; thus, loss aversion could be triggered and exhibited even in gain domain decision problems.[43]

Further, the estimates of the loss aversion parameter λ are between 2 and 4 in the four specifications. These estimates are similar in magnitude to those obtained in Chapter 1, as well as those found by Harrison and Rutström (2008) using similar specifications. Using Wald tests, I find that these estimates are statistically distinguishable from 1 at the 1% level in all four cases. Thus, the hypothesis of no loss aversion ($\lambda = 1$) is rejected in favour of $\lambda > 1$, and there is evidence of considerable loss aversion in the decision-making process of participants. In other words, losses loom larger than gains in decision-makers' valuations of risky lotteries, where losses and gains are evaluated in relation to

[42] $\chi = 0$ would imply that participants determine the gain and loss frames directly from the sign of the lottery prize (Andersen et al. 2006a).

[43] This reference point could be "homegrown", that is, not related to the framing of the lottery but based on prior experiences or beliefs about the experimental task (Andersen et al. 2006a, Crosetto and Filippin 2013). For decision problems in the loss domain or Binswanger-style lotteries in the gain domain, the initial endowment and the amount in the certain (risk-free) choice option, respectively, represent likely reference points for decision-making that are determined by the framing of the problems. However, in the case of gain domain MPL decision problems (as is the case here), it is difficult to assess *a priori* what, if any, are the likely reference points determined by framing. One candidate for a reference point determined by the framing of the problem is 0, since this is the status quo in a pure gain problem (Munro and Sugden 2003) – the estimated reference points in all four specifications in Table 6 are statistically significantly different from 0, providing some indication that the reference point may be "homegrown". However, given heterogeneity of subjects' unobserved experiences, it is possible that "homegrown" reference points just look like errors in variables for the estimated reference point.

the estimated reference point. Numerous studies have found that individuals exhibit loss aversion in both experimental behaviour (for example, Humphrey and Verschoor 2004a, Barr and Packard 2005) as well as non-experimental behaviour (for example, Farber 2008, Crawford and Meng 2011); additionally, Rabin (2000) notes that preferences incorporating loss aversion, unlike expected utility models, can explain both non-trivial risk aversion over small stakes and reasonable degrees of risk aversion over large stakes. Further, Fafchamps (2009) and Tanaka and Munro (2012) note that loss aversion may be a more important characterization of behaviour in developing countries than risk aversion, as it plays a larger role in shaping individual preferences and driving decision-making.[44]

In the specifications without the noise parameter, α_P is estimated to be 0.67 for Latent Index 1 and 0.54 for Latent Index 2. Since the estimated $\alpha_P < 1$, it implies that the utility function is concave over gains and convex over losses; that is, participants are risk averse over gains and risk seeking over losses, relative to the reference point of approximately R$4. Thus for these two specifications, the estimates of χ, λ and α_P imply that the utility function has a "S shape", as illustrated in Figure 1 of Chapter 1, and is steeper for losses than for gains near the reference point of R$4. These estimates are similar to those obtained in Chapter 1 for the two-part power utility specification.

On the other hand, the estimates of α_P are around 1.6 in both the error specifications. Further, these estimates are statistically significantly different from 0 at the 1% level. Additionally, the estimated noise parameter μ is approximately 11 for the Luce error specification and 16 for the Fechner error specification. These large (relative to 0) and statistically significant estimates of the noise parameter indicate that observed choices may deviate significantly

[44] Segal and Spivak (1990) define rank-dependent and loss aversion models as exhibiting first-order risk aversion, while expected utility models exhibit second-order risk aversion – first (second) order risk aversion implies that the risk premium for a small risk is proportional to the standard deviation (variance) of the risk. Further, Schmidt (1999) notes that the two types of risk aversion generate significantly different results, and models exhibiting first-order risk aversion can better explain real-world data and phenomena (for example, the equity premium puzzle).

from the predictions of the two-part power utility model, and there are significant stochastic errors (or randomness) in experimental decision-making.

So far, the decision models analyzed in this chapter use objective probabilities to weight the utilities from different prize outcomes in a lottery. However, the remaining decision models considered in this section allow for subjective probability weighting. These models fall under the umbrella of rank-dependent utility theory – developed by Quiggin (1982, 1993) – which does not assume that the utilities from different outcomes in a lottery are weighted by objective probabilities, but instead allows decision-makers to use non-linear transformations of cumulative objective probabilities for weighting outcomes. Thus, RDU allows preferences to depend on the rank of the final outcome through probability weighting, and the probability weighting function can be interpreted as a decision weighting function which measures the subjective importance of different outcomes in the evaluation of a lottery (Humphrey and Verschoor 2004a, Harrison and Rutström 2008).

There are two components in the RDU specification: the utility function and the probability weighting function. In the RDU framework, probability weighting can be combined with any of the three utility functions described earlier (CRRA, EP and two-part power utility). However, instead of weighting outcomes with the associated objective probabilities p_{k_i}, RDU uses transformations of cumulative probabilities, w_{k_i}, to weight outcomes and evaluate lotteries, as follows:

$$RDU_i = \sum_{k_i=1}^{K} w_{k_i} U_{k_i} \tag{15}$$

where RDU_i is the rank-dependent utility from lottery i and the states (outcomes) in lottery i are ranked from worst (U_{1_i}) to best (U_{K_i}). w_{k_i} is defined as follows:

$$w_{k_i} = \begin{cases} \omega(p_{1_i}) & \text{for } k_i = 1 \\ \omega(p_{1_i} + \ldots + p_{k_i}) - \omega(p_{1_i} + \ldots + p_{(k-1)_i}) & \text{for } k_i > 1 \end{cases} \tag{16}$$

where $\omega(p)$ is some probability weighting function.

As is clear from Equations (15) and (16), the probability weighting function (ω) is an important component of RDU specifications. In this chapter, I consider the widely used weighting function proposed by Tversky and Kahneman (1992), which is given by[45]:

$$\text{Tversky-Kahneman function: } \omega^{TK}(p) = p^{\gamma}/[p^{\gamma} + (1-p)^{\gamma}]^{1/\gamma} \tag{17}$$

for $0 < p < 1$. Weighting of cumulative objective probabilities is implied by $\gamma \neq 1$, and this function has well-defined endpoints, with $\omega^{TK}(p) = 0$ for $p = 0$ and $\omega^{TK}(p) = 1$ for $p = 1$.

Thus, in the RDU models, outcomes are aggregated using Equations (15), (16) and (17); this probability weighting procedure is combined with the CRRA and EP utility functions in order to obtain log-likelihood functions, using Latent Index 1 and 2 (both with and without the noise parameter). These log-likelihood functions can then be optimized to obtain structural maximum likelihood estimates of the utility function parameters and the probability weighting parameter γ (as well as the noise parameter μ in the error specifications). The results of the RDU-CRRA estimations are provided in Table 7 and the results of the RDU-EP estimations are presented in Table 8.

The estimates of the coefficient of relative risk aversion r in the CRRA utility function are between 0.3 and 0.6 for the four RDU-CRRA specifications in Table 7. These estimates are statistically significantly different from 0 at the 1% level, indicating that participants exhibit moderate risk aversion in the experimental decision problems.

The estimates of the r_{EP} parameter in the four RDU-EP specifications (Table

[45] I also conduct all the RDU estimations using the Prelec (1998) probability weighting function, which is another weighting function that has been used extensively in analyses of RDU models. However, the results in all cases are very similar to those obtained using the Tversky and Kahneman (1992) function, and so I only report the estimation results for the specifications using the Tversky-Kahneman probability weighting function for conciseness.

Table 7. *Structural maximum likelihood estimates of risk attitudes for RDU with CRRA utility and the Tversky-Kahneman probability weighting function*

Coefficient	Estimate	Standard Error	p-value	Lower 95% Confidence Interval	Upper 95% Confidence Interval	Observations
A. Latent Index 1: no error parameter						
r	0.342	0.012	0.000	0.319	0.365	7452
γ	0.670	0.018	0.000	0.635	0.705	7452
$H_0 : \gamma = 1$			0.000			
B. Latent Index 1: Luce error parameter						
r	0.459	0.060	0.000	0.342	0.575	7452
γ	0.687	0.017	0.000	0.654	0.720	7452
μ	0.760	0.103	0.000	0.558	0.963	7452
$H_0 : \gamma = 1$			0.000			
C. Latent Index 2: no error parameter						
r	0.562	0.011	0.000	0.540	0.585	7452
γ	0.682	0.019	0.000	0.645	0.719	7452
$H_0 : \gamma = 1$			0.000			
D. Latent Index 2: Fechner error parameter						
r	0.424	0.056	0.000	0.315	0.533	7452
γ	0.689	0.018	0.000	0.653	0.724	7452
μ	1.378	0.174	0.000	1.036	1.720	7452
$H_0 : \gamma = 1$			0.000			

The log-likelihoods corresponding to the estimations in Panels A, B, C and D are -4253.4, -4251.8, -4260.9 and -4256.6.
Wald test p-values and 95% confidence intervals reported.

Table 8. *Structural maximum likelihood estimates of risk attitudes for RDU with expo-power utility and the Tversky-Kahneman probability weighting function*

Coefficient	Estimate	Standard Error	p-value	Lower 95% Confidence Interval	Upper 95% Confidence Interval	Observations
A. Latent Index 1: no error parameter						
r_{EP}	0.136	0.056	0.016	0.026	0.246	7452
α_{EP}	0.041	0.018	0.027	0.005	0.077	7452
γ	0.684	0.023	0.000	0.639	0.729	7452
$H_0 : \gamma = 1$			0.000			
B. Latent Index 1: Luce error parameter						
r_{EP}	0.525	0.146	0.000	0.240	0.810	7452
α_{EP}	-0.074	0.190	0.698	-0.445	0.298	7452
γ	0.692	0.021	0.000	0.651	0.733	7452
μ	0.389	0.095	0.000	0.202	0.576	7452
$H_0 : \gamma = 1$			0.000			
C. Latent Index 2: no error parameter						
r_{EP}	0.346	0.072	0.000	0.205	0.487	7452
α_{EP}	0.016	0.055	0.777	-0.092	0.123	7452
γ	0.690	0.022	0.000	0.647	0.732	7452
$H_0 : \gamma = 1$			0.000			
D. Latent Index 2: Fechner error parameter						
r_{EP}	0.472	0.142	0.001	0.194	0.750	7452
α_{EP}	-0.043	0.140	0.760	-0.317	0.231	7452
γ	0.693	0.021	0.000	0.652	0.734	7452
μ	0.757	0.182	0.000	0.401	1.114	7452
$H_0 : \gamma = 1$			0.000			

The log-likelihoods corresponding to the estimations in Panels A, B, C and D are -4256.0, -4251.6, -4257.0 and -4256.5.
Wald test p-values and 95% confidence intervals reported.

8) are positive; additionally, the hypothesis of constant absolute risk aversion ($r_{EP} = 0$) is rejected in all four cases (at the 1% or 5% significance levels) in favour of decreasing absolute risk aversion ($r_{EP} > 0$). The estimates of α_{EP} are extremely small in magnitude – in all specifications except that presented in Panel A, the hypothesis of CRRA ($\alpha_{EP} = 0$) cannot be rejected at the 10% significance level.

The estimates of the probability weighting parameter are around 0.69 in all the RDU-CRRA and RDU-EP specifications. These estimates are statistically significantly different from 1 – thus, using Wald tests, the null hypothesis of no probability weighting ($\gamma = 1$) is rejected at the 1% level in all eight specifications in favour of $\gamma < 1$. It is important to note that the RDU-CRRA and RDU-EP specifications nest the EUT-CRRA and EUT-EP models, respectively; that is, when $\gamma = 1$, subjects weight outcomes by the associated objective probabilities, and the RDU specifications collapse to the corresponding EUT specifications. Thus, a rejection of $\gamma = 1$ implies a rejection of EUT as the latent decision-making process of participants. Further, since $0 < \gamma < 1$, there is evidence that probability weighting follows an "inverse S-shape" (as shown in Figure 2 of Chapter 1), concave for low probabilities and convex for high probabilities (Gonzalez and Wu 1999). Additionally, since probability weights can be interpreted as decision weights in the RDU setting, these parameter values indicate an overweighting of extreme (unlikely) outcomes (Harrison and Rutström 2008). There is substantial evidence of "inverse S-shaped" probability weighting in the experimental literature (Gonzalez and Wu 1999), and these estimates of γ are very similar to those obtained in numerous experimental studies (for example, Tversky and Kahneman 1992, Camerer and Ho 1994, Harrison et al. 2007). However, some studies conducted in developing countries, including the analysis in Chapter 1, have instead found evidence of "S-shaped" probability weighting (for example, Humphrey and Verschoor 2004b, Harrison et al. 2010)

In all but one of the four RDU error specifications (Tables 7 and 8), μ is estimated to be less than 1. While the hypothesis of no stochastic errors ($\mu = 0$) is rejected at the 1% level in all four error specifications, the estimates of the noise parameter relatively close to 0 indicate that if RDU (combined with the CRRA or EP utility function) is assumed to represent the true decision-making process of participants, there are expected to be relatively few deviations in observed choices from true preferences.

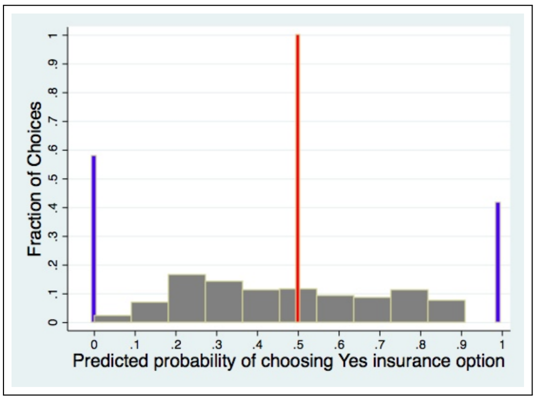
To illustrate the extent of randomness in subject choice implied by the noise parameter estimates, I construct histograms displaying individuals' predicted choice probabilities of the Yes insurance choice for each of the RDU and EUT error specifications. Figure 1 presents the histograms for the EUT-CRRA, EUT-EP, RDU-CRRA and RDU-EP models combined with the Luce error specification, while Figure 2 presents the corresponding histograms for the Fechner error specification. For each histogram, the maximum likelihood estimates of the preference and noise parameters are used to construct the displayed predicted probabilities[46]; in addition, predicted probabilities in the case of deterministic choice ($\mu \rightarrow 0$) – where observed choices only depend on the utilities obtained from the Yes and No insurance options – and completely random choice ($\mu \rightarrow \infty$) – where observed choices do not depend on underlying preferences and utilities – are displayed in each of the histograms.

The figures indicate that there is considerable randomness in choice for all of the decision models, error specifications and noise parameter μ estimates. The histograms for the RDU specifications, which involve notably smaller noise parameter estimates than the corresponding EUT specifications, show a similar distribution of predicted probabilities as compared to the histograms for the EUT specifications. However, the figures indicate that in the RDU specifications, the fractions of choices in the two extreme predicted probability categories (that is, predicted probabilities closest to 0 and 1) are marginally greater than in the corresponding EUT specifications. All the model estimations indicate a relatively large fraction of choices for which the predicted probability

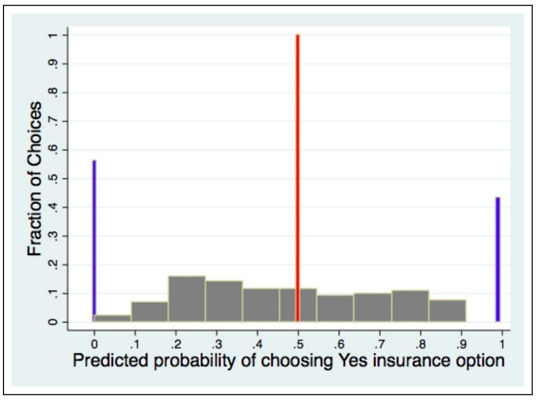
[46] Note that choice data from all the decision problems (and all questions) are pooled in the estimation of the noise and preference parameters, as well as in the construction of the histograms.

Figure 1. *Extent of randomness in choice: Luce error parameter*

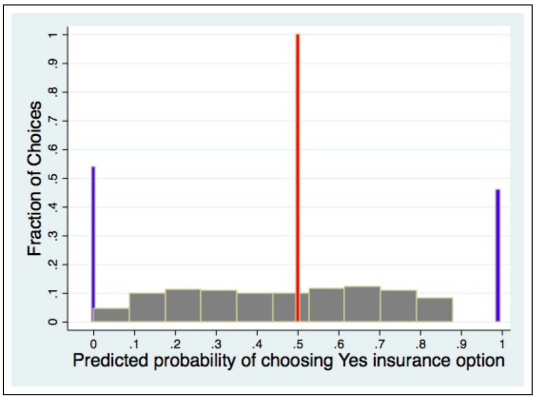
(a) EUT-CRRA: Estimated $\mu = 72.64$



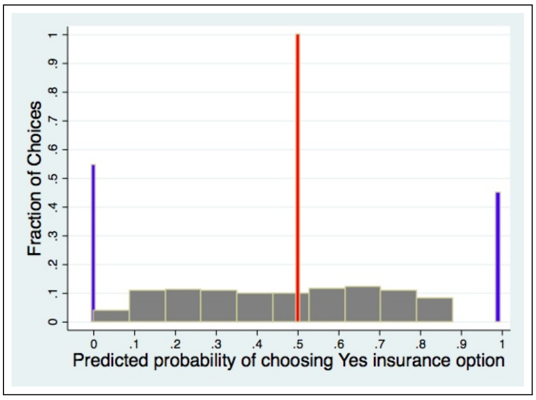
(b) EUT-EP: Estimated $\mu = 5.063$



(c) RDU-CRRA: Estimated $\mu = 0.760$



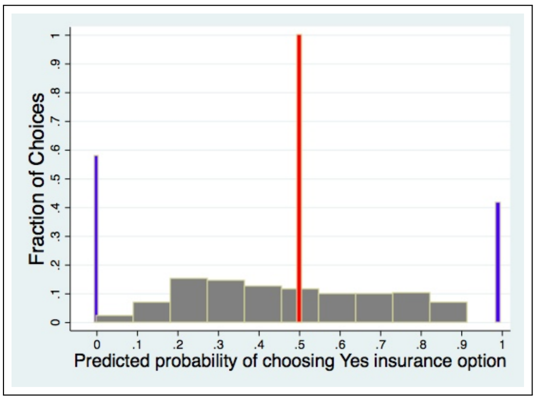
(d) RDU-EP: Estimated $\mu = 0.389$



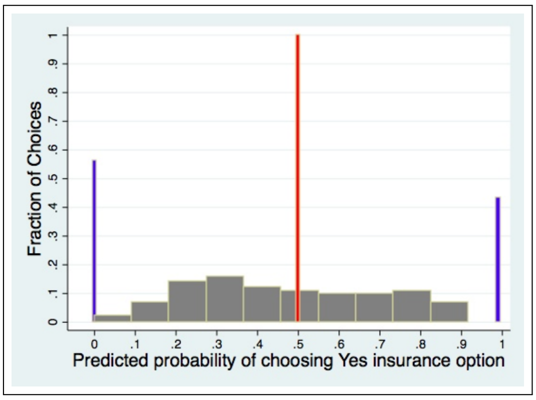
Red: $\mu \rightarrow \infty$; Blue: $\mu \rightarrow 0$; Gray: Maximum Likelihood Estimate of μ

Figure 2. *Extent of randomness in choice: Fechner error parameter*

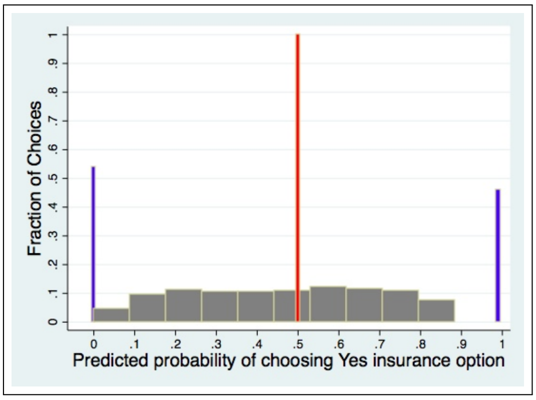
(a) EUT-CRRA: Estimated $\mu = 117.6$



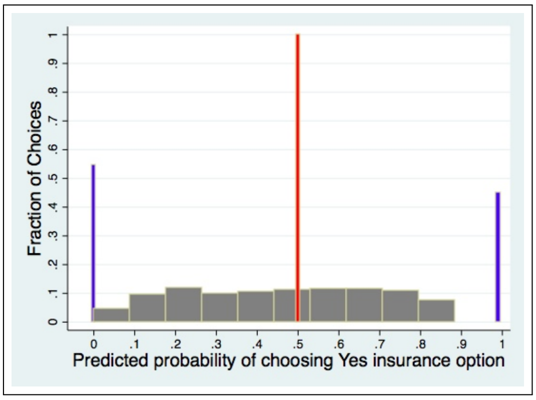
(b) EUT-EP: Estimated $\mu = 8.222$



(c) RDU-CRRA: Estimated $\mu = 1.378$



(d) RDU-EP: Estimated $\mu = 0.757$



Red: $\mu \to \infty$; Blue: $\mu \to 0$; Gray: Maximum Likelihood Estimate of μ

(implied by the structural model) is close to 0.5 (which implies randomizing with equal probability between the Yes and No alternatives). This suggests that there is significant randomness in choice no matter which decision model is assumed to be the true decision-making process – even though the estimated μ parameters are notably different in the different models, each of the estimated noise parameter values implies substantial deviation in observed choices from true preferences, with only a small fraction of participants choosing the Yes or No option with near certainty depending on their true preferences. This highlights the possibility that none of the standard models considered provide a suitable (or accurate) description of the true decision-making process of participants; this is investigated further in Section 4.3.

Additionally, this provides an indication that the magnitude of the noise parameter estimate does not depend solely on the extent of randomness in choice, but also depends on the preference parameters included in the model (and the estimates of these parameters). In other words, the finding that substantially different estimates of μ imply similar (high) levels of randomness in choice indicates that, as noted earlier, a comparison of the extent of noise in decision-making cannot be conducted through a direct comparison of μ (across contexts in which the preference parameters differ); such a comparison does not have a cardinal or ordinal interpretation.

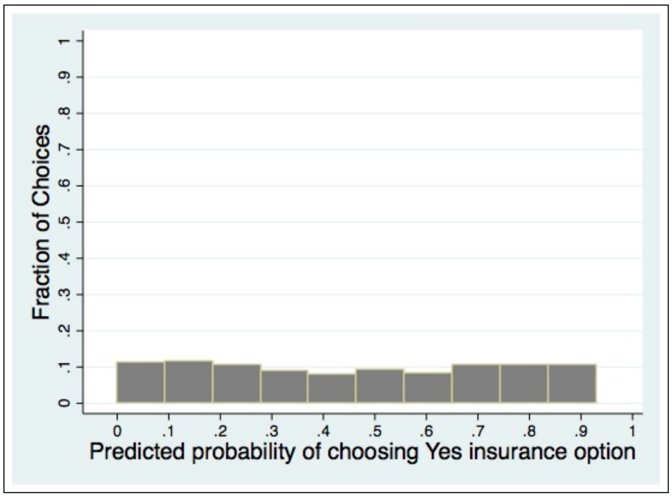
To further explore the relationship between the magnitude of the noise parameter and the extent of randomness in observed choice, I construct similar histograms for lower (hypothetical) values of μ, assuming that RDU-EP (which has the lowest μ estimate of all the RDU and EUT specifications) reflects the true decision-making process of participants, with preference parameters set equal to the values obtained in the earlier maximum likelihood estimations. The histograms displaying choice probabilities corresponding to μ values of 0.05, 0.1 and 0.3 for the Luce error specification are presented in Figure 3, and the histograms for the Fechner error specification and the same μ values are displayed in Figure 4. These figures indicate that given RDU-EP preferences with preference parameter values as provided by the earlier maximum likelihood estimations, there is still considerable randomness in choice when $\mu = 0.3$. However, when μ falls to 0.1 and 0.05 (that is, even closer to the no error value of 0), there is substantially less deviation in observed choices from the predictions of the RDU-EP model – a large fraction of subjects have predicted probabilities of choosing the Yes insurance option that are close to 0 or 1; thus, a large fraction of choices follow true preferences with near certainty, with less randomness in choice (and there are only a small fraction of choices with predicted probabilities close to 0.5).

To shed light on the suitability of each of the estimated EUT and RDU error specifications for predicting observed choices, I constructed graphs showing the predicted probability (by question) of choosing the Yes insurance option (indicated by the structural model) and the observed fraction (by question) of Yes choices, for each of the 20 questions (averaged across the seven decision problems). The graphs for the Luce error specification are presented in Figure 5 and those for the Fechner error specification are displayed in Figure 6. These graphs indicate that, for each decision model and error specification, there is substantial difference in the probability of Yes choice predicted by the structural model and the observed fraction of Yes choices, for most of the 20 questions (averaged across the seven decision problems). This reflects the deviation in observed choice fractions from the predictions of the EUT and RDU models, and once again highlights the possibility that even when the extent of errors in decision-making is accounted for, none of these decision models provide an accurate description of the true decision-making process of subjects. Additionally, the magnitudes of these deviations are similar for the different decision models, which involve notably different estimates of μ.

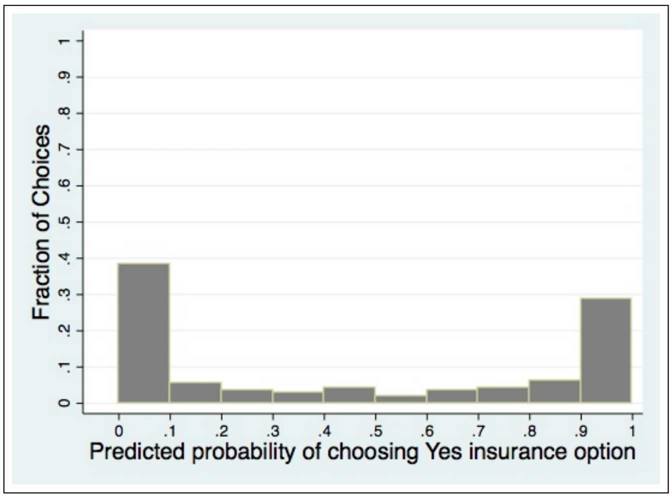
The graphs also indicate that, as expected, at lower prices (that is, in earlier questions) the observed fraction of choices indicating preference for buying the insurance contract is high (nearly 85%), and this fraction decreases as the price of the insurance contract increases (that is, as the question number increases). The

Figure 3. *Extent of randomness in choice with hypothetical μ values: Luce error parameter*

(a) RDU-EP: $\mu = 0.3$



(b) RDU-EP: $\mu = 0.1$



(c) RDU-EP: $\mu = 0.05$

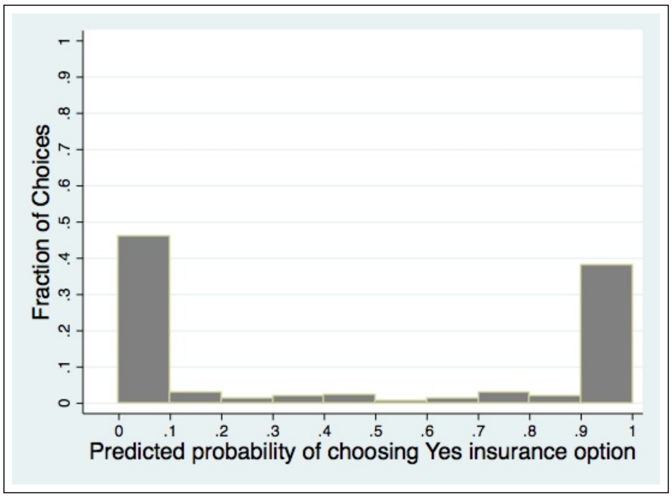
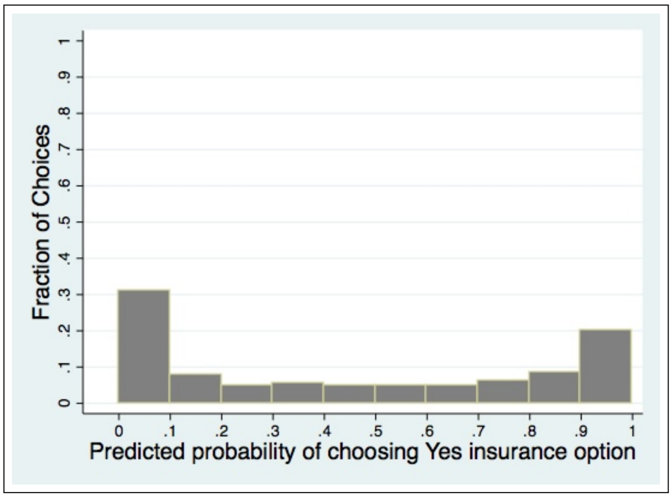
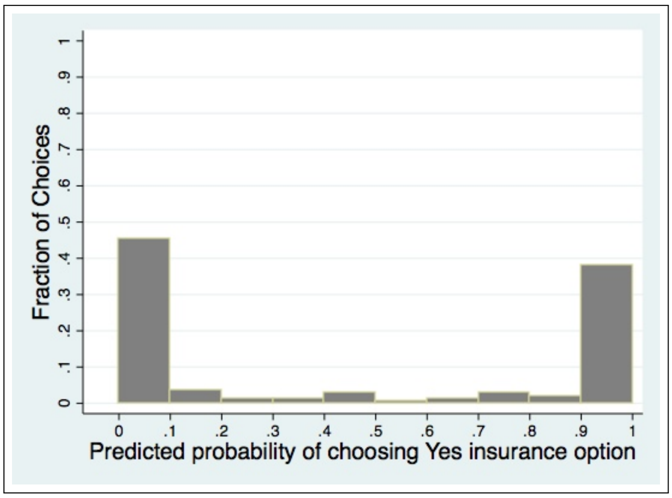


Figure 4. *Extent of randomness in choice with hypothetical μ values: Fechner error parameter*

(a) RDU-EP: $\mu = 0.3$



(b) RDU-EP: $\mu = 0.1$



(c) RDU-EP: $\mu = 0.05$

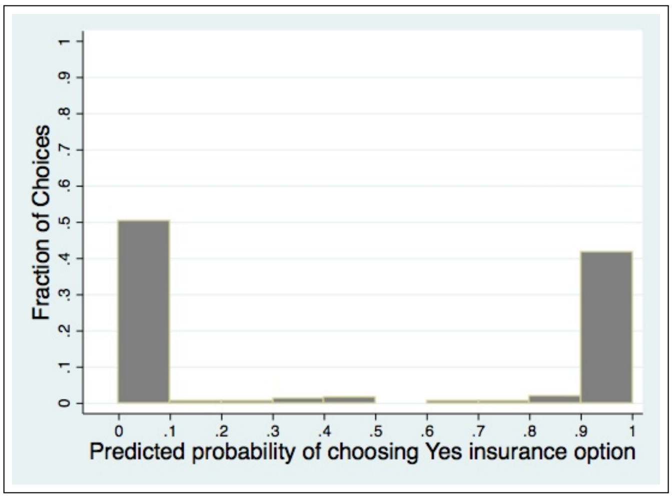
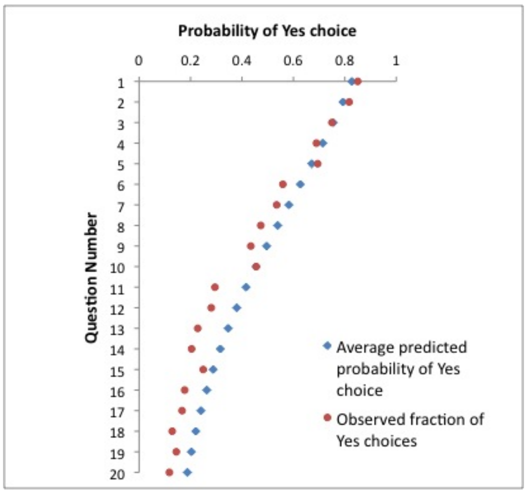
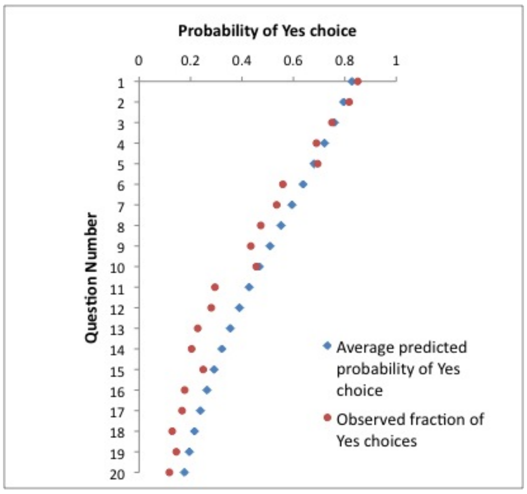


Figure 5. *Average predicted probability of Yes insurance choice by question: Luce error parameter*

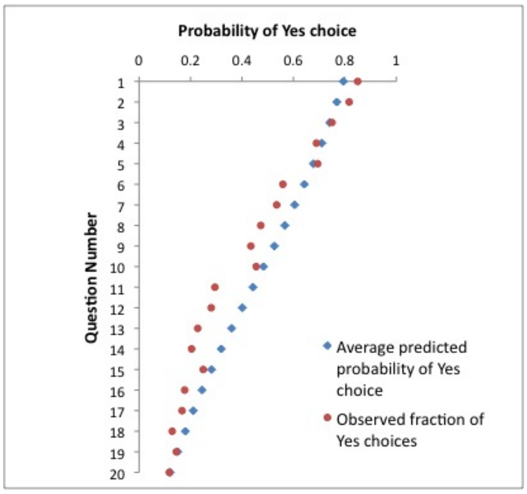
(a) EUT-CRRA: Estimated $\mu = 72.64$



(b) EUT-EP: Estimated $\mu = 5.063$



(c) RDU-CRRA: Estimated $\mu = 0.760$



(d) RDU-EP: Estimated $\mu = 0.389$

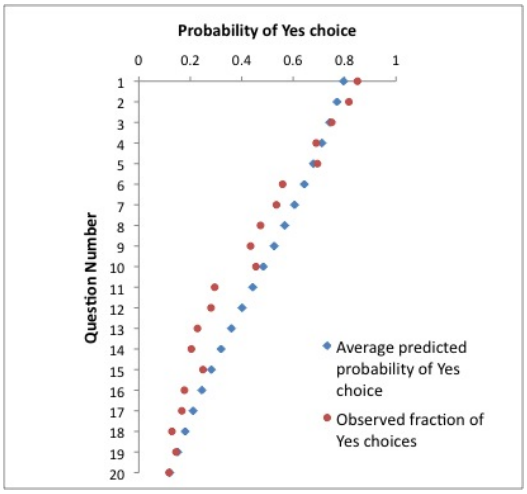
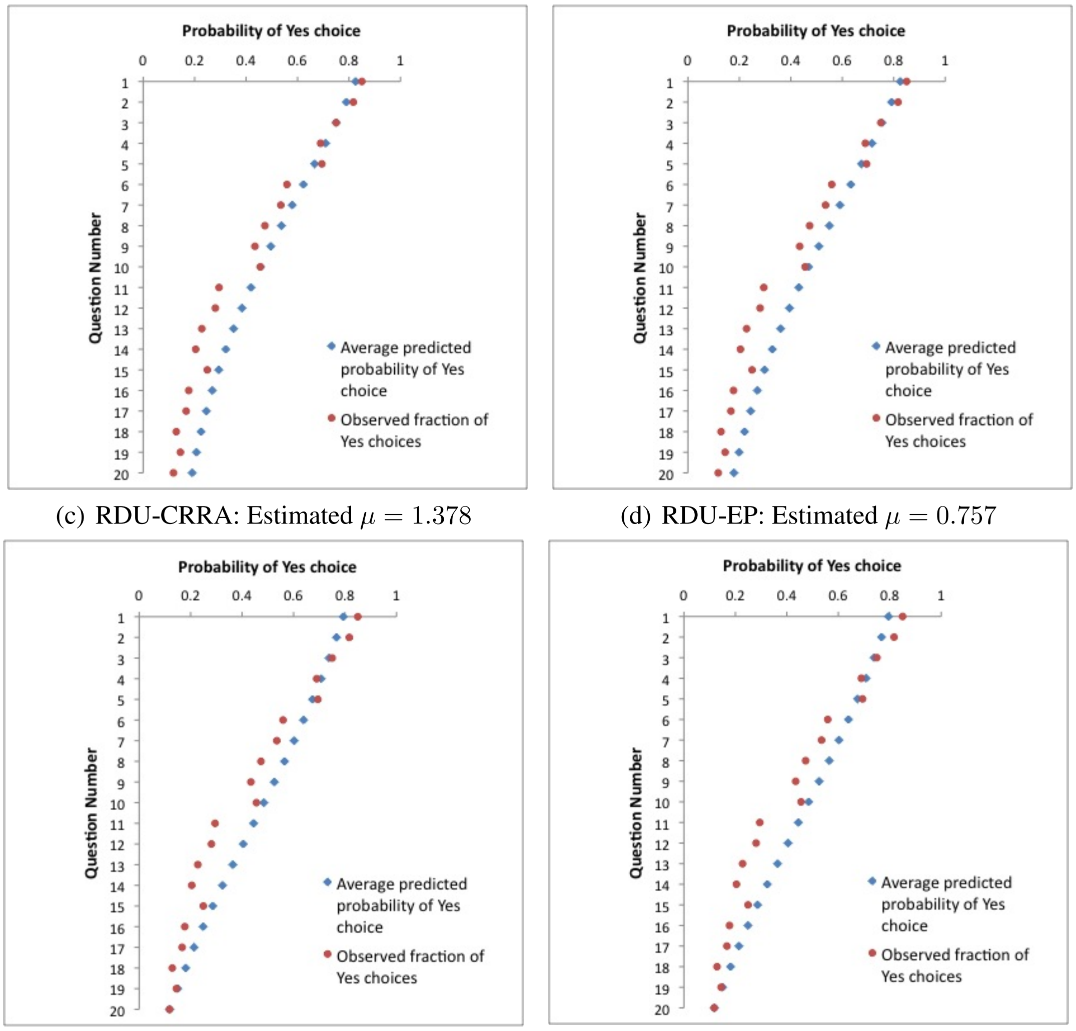


Figure 6. *Average predicted probability of Yes insurance choice by question: Fechner error parameter*



(a) EUT-CRRA: Estimated $\mu = 117.6$

(b) EUT-EP: Estimated $\mu = 8.222$

(c) RDU-CRRA: Estimated $\mu = 1.378$

(d) RDU-EP: Estimated $\mu = 0.757$

observed fraction of Yes insurance choices is only 12% in question 20, when the price of the insurance contract is highest (R$8).[47]

Rank-dependent utility theory can also be extended to utility functions which are defined separately over gains and losses relative to a reference point (rather than final outcomes), such as the two-part power utility function given in Equation (14). This describes a class of decision models comprising cumulative prospect theory (CPT), which was developed by Tversky and Kahneman (1992) and has emerged as one of the main contenders to the conventional EUT model as a descriptor of decision-making under risk (Starmer 2000). Thus, the final specification estimated in this section is the CPT model, which uses the Tversky-Kahneman weighting function (Equation (17)) to weight probabilities, rank-dependent aggregation of outcomes in a lottery (Equations (15) and (16)), and the two-part power utility function (Equation (14)) to evaluate each outcome.[48] This specification is commonly used in studies involving CPT (for example, Liu 2008, Tanaka et al. 2010). The latent indices are constructed as before (both with and without the noise parameter), and the corresponding log-likelihood functions can then be maximized to provide structural maximum likelihood estimates of the various parameters. The results of these CPT estimations are provided in Table 9.

The results in Table 9 once again indicate that the estimated reference point is around R$4 in the specifications without the noise parameter, and R$8 in the specifications with the noise parameter. As found in the error specifications for

[47] However, this pattern of the observed fraction of “safe” choices differs significantly from that obtained by Holt and Laury (2002), who analyze MPL data from the United States and find that nearly 100% of participants choose the “safe” option in the first question and almost none of the participants choose the “safe” option in the final question. On the other hand, Charness and Viceisza (2012), conduct an amended version of the Holt and Laury (2002) decision problem in rural Senegal and observe relative insensitivity in observed choices to the question number. They find that only 60% of subjects choose the “safe” option in the first question and 40% of subjects still choose the “safe” option in the final question; further, the observed fraction of “safe” choices does not decrease monotonically from the first to the last question. They interpret this as an indication of within-subject inconsistency in choices and a lack of participant understanding of the MPL decision problem.

[48] Since the RDU-CRRA and RDU-EP specifications defined in this chapter do not allow for differential evaluation of gains and losses (relative to a possibly non-zero reference point), which is a critical aspect of cumulative prospect theory, I do not refer to them as CPT specifications. Chapter 1 uses a similar classification.

Table 9. *Structural maximum likelihood estimates of risk attitudes for CPT with the Tversky-Kahneman probability weighting function*

Coefficient	Estimate	Standard Error	p-value	Lower 95% Confidence Interval	Upper 95% Confidence Interval	Observations
A. Latent Index 1: no error parameter						
χ	4.278	0.367	0.000	3.559	4.997	7452
λ	2.031	0.235	0.000	1.569	2.493	7452
α_P	0.706	0.011	0.000	0.685	0.727	7452
γ	0.710	0.022	0.000	0.667	0.753	7452
$H_0 : \lambda = 1$			0.000			
$H_0 : \gamma = 1$			0.000			
B. Latent Index 1: Luce error parameter						
χ	8.002	0.076	0.000	7.853	8.151	7452
λ	3.187	0.326	0.000	2.548	3.826	7452
α_P	1.524	0.068	0.000	1.390	1.658	7452
γ	0.914	0.029	0.000	0.857	0.971	7452
μ	8.116	1.329	0.000	5.512	10.720	7452
$H_0 : \lambda = 1$			0.000			
$H_0 : \gamma = 1$			0.005			
C. Latent Index 2: no error parameter						
χ	3.722	0.165	0.000	3.398	4.046	7452
λ	0.891	0.157	0.000	0.584	1.198	7452
α_P	0.574	0.009	0.000	0.556	0.593	7452
γ	0.689	0.020	0.000	0.649	0.729	7452
$H_0 : \lambda = 1$			0.486			
$H_0 : \gamma = 1$			0.000			
D. Latent Index 2: Fechner error parameter						
χ	8.020	0.364	0.000	7.306	8.734	7452
λ	3.038	0.630	0.000	1.803	4.274	7452
α_P	1.522	0.100	0.000	1.327	1.718	7452
γ	0.910	0.030	0.000	0.851	0.969	7452
μ	13.44	3.611	0.000	6.360	20.52	7452
$H_0 : \lambda = 1$			0.000			
$H_0 : \gamma = 1$			0.003			

The log-likelihoods corresponding to the estimations in Panels A, B, C and D are -4224.6, -4162.3, -4233.7 and -4167.5.
Wald test p-values and 95% confidence intervals reported.

the decision models considered earlier, the estimates of the noise parameter (μ) are statistically significantly different from both 0 and 1 (Panels B and D).[49] Additionally, in the two specifications involving Latent Index 1, as well as the Latent Index 2 specification with the error parameter, the hypothesis of $\lambda = 1$ is rejected at the 1% significance level in favour of $\lambda > 1$ – thus, as found for the two-part power utility specification without probability weighting, there is evidence of loss aversion in subjects' decision-making.

The estimates of the curvature parameter α_P are also similar to those reported in Table 6 (which presents the estimation results for the PU specification without probability weighting). In addition, the estimates of γ in all four specifications are less than 1, indicating "inverse S-shaped" probability weighting; these estimates are statistically distinguishable from 1 at the 1% level.[50,51]

4.2 Correlates of stochastic errors and inconsistencies in choice

As detailed in the previous section, using the MPL data from the Brazilian experiment, a larger number of decision models and preference parameters can be estimated using the structural maximum likelihood procedure, as compared to when using the Ordered Lottery Selection data from the Ethiopian experiment (which includes a similar number of participants and decision problems). Dave et al. (2010), however, note that the additional complexity and precision of

[49] Further, I constructed graphs illustrating the extent of randomness in choice implied by the noise parameter estimates for the CPT and two-part power utility error models, similar to the graphs presented in Figures 1 and 2. The extent of randomness is very similar to that indicated by the EUT and RDU error estimations, as displayed in Figures 1 and 2. I also constructed graphs illustrating the deviation in observed choices from the predicted choice probabilities (by question) implied by the CPT and two-part power utility error estimations, obtaining similar representations as those in Figures 5 and 6 for the EUT and RDU error estimations.

[50] It is important to point out that, for most of the specifications reported in this section, the estimates obtained by both latent indices are similar, providing some indication that the two latent indices do not produce markedly different characterizations of risk preferences in binary choice experiments (which has also been noted by Harrison and Rutström 2008).

[51] It is also important to note that all the parameters of the CPT model, as well as of the RDU and two-part power utility models, are estimated (with relatively narrow confidence intervals) using data from the seven decision problems of the Brazilian experiment, whereas I encountered non-convergence of the optimization procedure in the corresponding maximum likelihood estimations using the Ordered Lottery Selection data from the Ethiopian experiment (see Chapter 1).

the MPL procedure come at a cost. Using data from an experiment – which included both MPL and Ordered Lottery Selection decision problems – conducted with 900 adults in Canada, they study the tradeoffs between the two types of elicitation procedures and find that subjects with low math ability and lower education levels, in particular, face much more difficulty in understanding the MPL procedure as compared to the Ordered Lottery Selection method, leading to substantially more randomness (or noise) observed in their decisions in the MPL problem.[52] Similarly, Crosetto and Filippin (2013) also find evidence of greater noise in subjects' decisions in MPL decision problems as compared to Ordered Lottery Selection problems; they attribute this to the greater "cognitive load" placed on participants by the more complex MPL procedure, which hinders comprehension. However, appropriate understanding of decision problems is crucial for eliciting meaningful responses that reflect the true risk attitudes of subjects (Charness et al. 2013).

To explore the possibility of poor comprehension – and to shed light on the correlates of stochastic errors in experimental decision-making – I extend the analysis conducted in the previous section to account for possible heterogeneity in the noise parameter. In this section, I estimate the Luce and Fechner error specifications in combination with the EUT-CRRA, EUT-EP, PU, RDU-CRRA, RDU-EP and CPT decision models, but no longer assume the noise parameter (μ) to be constant across participants. Instead the noise parameter is considered to be a linear function of various individual characteristics of the subject, while the other (preference) parameters in the decision models are assumed to be constant across subjects.

The estimation strategy utilized is similar to that used in Chapter 2 to analyze the determinants of risk preferences (described in Section 2.3 of Chapter 2), and has been used widely in the experimental literature (for example, Harrison and Rutström 2008, Harrison et al. 2007, Dave et al. 2010). It is also similar to

[52] They also find that subjects in their experiment requested greater clarification from enumerators for the MPL decision problem as compared to the Ordered Lottery Selection problem.

the structural maximum likelihood procedure detailed in the previous section; the only difference is that in this case, the noise parameter is not assumed to be constant across subjects, but is allowed to vary by individual characteristics. Thus, in this empirical analysis, for each participant a, $\mu_a = \mu(\boldsymbol{X})$, where $\boldsymbol{X}$ is a vector consisting of four individual characteristics – age, gender, education and wealth. These variables, described in Section 3.2, are along the lines of those used by Dave et al. (2010) in similar estimations, and are *a priori* expected to be important correlates of stochastic errors in experimental decision-making. Following Harrison et al. (2007) and Harrison and Rutström (2008), μ_a can therefore be described by:

$$\mu_a = \boldsymbol{\beta X_a} \tag{18}$$

$$\mu_a = \beta_0 + \beta_1 Age_a + \beta_2 Gender_a + \beta_3 Education_a + \beta_4 Wealth_a \tag{19}$$

where the education variable is a binary dummy indicating whether the subject had completed – or was currently enrolled in – college at the time of the experiment, and the wealth variable is the PCA wealth index described in Section 3.2.

Thus, the estimation of μ_a entails the estimation of the five coefficients comprising the vector $\boldsymbol{\beta}$. The latent indices and the log-likelihoods for the different decision models are constructed in the same way as described in Section 4.1. However, since the noise parameter μ is a linear function of the individual characteristics comprising $\boldsymbol{X}$, the log-likelihoods are maximized with respect to the preference parameters in the decision models (that is, all parameters other than the noise parameter) as well as the five β coefficients. In this way, I obtain maximum likelihood estimates of the preference parameters, which are assumed to be constant, and the five β coefficients (which determine the estimate of μ). The results of these estimations for the different decision models in conjunction with the Luce error specification are presented in Table 10, while Table 11

Table 10. *Correlates of the Luce error parameter*

Variables	EUT-CRRA (1)	EUT-EP (2)	PU (3)	RDU-CRRA (4)	RDU-EP (5)
Age	3.892**	0.451	0.180**	0.000373	0.000572
	(1.661)	(0.769)	(0.0770)	(0.00838)	(0.00527)
Gender	-2.009	-0.295	-0.808*	-0.137**	-0.0819*
	(4.843)	(0.750)	(0.470)	(0.0570)	(0.0444)
College education	-52.00**	-6.111	-6.456***	-0.774***	-0.469**
	(22.14)	(10.33)	(1.443)	(0.180)	(0.185)
PCA wealth	2.107	0.184	0.174	-0.00395	-0.00281
	(3.481)	(0.534)	(0.333)	(0.0407)	(0.0249)
Constant	36.66	4.931	9.755***	1.657***	1.000***
	(27.55)	(8.483)	(2.712)	(0.357)	(0.388)
Observations	7412	7412	7412	7412	7412
Log-likelihood	-4190.0	-4190.5	-4064.4	-4166.9	-4166.9

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$
Dependent variable is the noise parameter (μ) of the Luce error model.
Standard errors are in parentheses.

presents the results for the corresponding Fechner error estimations.[53]

However, the CPT model combined with the Luce or Fechner error specification could not be estimated due to non-convergence of the numerical optimization procedure in the maximum likelihood estimation, when μ is modeled as a linear function of the four individual characteristics.[54,55] Thus, Tables 10 and 11 only report the results for the estimation of the five other decision models – EUT-CRRA, EUT-EP, PU, RDU-CRRA and RDU-EP.

As shown in Tables 10 and 11, the coefficient estimates of age are positive in all

[53] Note that in these tables, I only present the estimates of the β coefficients for conciseness, since the estimates of the constant preference parameters are very similar to those reported in the previous section.

[54] This is the case when using the Newton-Raphson and Berndt-Hall-Hall-Hausman optimization algorithms in Stata as well as when using the sequential quadratic programming optimization algorithm in MATLAB.

[55] Further, for the maximum likelihood estimations in this chapter for which non-convergence of the optimization procedure was encountered, scans (as well as graphs of the log-likelihood function) indicate that the log-likelihood function is flat over a range of maximum likelihood parameter estimates – that is, the evidence indicates that multiple parameter estimates produce the same (maximized) value of the log-likelihood function, and there is insufficient information to identify unique maximum likelihood estimates. Similar results are obtained in Chapter 1.

Table 11. *Correlates of the Fechner error parameter*

Variables	EUT-CRRA (1)	EUT-EP (2)	PU (3)	RDU-CRRA (4)	RDU-EP (5)
Age	5.895**	0.525	0.288**	0.00312	0.00246
	(2.484)	(0.511)	(0.118)	(0.0147)	(0.0102)
Gender	-2.517	-0.323	-1.246*	-0.247**	-0.166*
	(6.902)	(0.708)	(0.693)	(0.0996)	(0.0907)
College education	-73.24**	-6.678	-10.00***	-1.334***	-0.901**
	(30.83)	(6.406)	(3.151)	(0.309)	(0.369)
PCA wealth	3.239	0.220	0.241	-0.0112	-0.00800
	(4.988)	(0.515)	(0.491)	(0.0715)	(0.0485)
Constant	51.39	5.556	15.90***	2.945***	1.987**
	(38.48)	(5.866)	(5.573)	(0.614)	(0.794)
Observations	7412	7412	7412	7412	7412
Log-likelihood	-4196.0	-4195.7	-4069.2	-4170.8	-4170.8

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$
Dependent variable is the noise parameter (μ) of the Fechner error model.
Standard errors are in parentheses.

the specifications (for both the Luce and Fechner error models).[56] Additionally, the estimates are statistically significantly different from 0 at the 5% level in the EUT-CRRA and PU models (for both the Luce and Fechner error specifications). Thus, these results provide some indication that older participants make greater stochastic errors in experimental decision-making. It is important to note that since the preference parameters are assumed to be constant across subjects, μ estimates can be directly compared in order to obtain a comparison of the extent of errors in decision-making across subjects with different individual characteristics (within a particular decision model).[57]

The coefficient estimates of gender are negative in all five columns of both tables; additionally, they are statistically significant (at the 5% or 10% level) in the PU, RDU-CRRA and RDU-EP specifications. These estimates indicate that women exhibit significantly lower noise in experimental decision-making. Further, there

[56] The age data is missing for two participants (see Table 3) – as a result, there are 40 fewer choices included in the estimations reported in this section than those reported in in Section 4.1.

[57] It is possible that risk preferences vary by individual characteristics and this interacts with the specification. However, if I allow preference parameters to vary across subjects (for example, by individual characteristics), I would not be able to compare the extent of noise in decision-making through a direct comparison of the noise parameter, which is the focus of this analysis.

is no evidence that wealth is related to the noise parameter – the coefficient estimates of the PCA wealth measure are statistically insignificant in all the specifications, for the both the Luce and Fechner error models. The coefficient estimates of the education variable are negative for all decision models, and are statistically significant (at the 1% or 5% level) for all models except the EUT-EP specification (Column (2)). This provides strong evidence that subjects who did not attend college exhibit significantly greater noise in their decisions in the MPL problems as compared to those who attended college (assuming the same risk preferences across both groups), in line with the results of Dave et al. (2010).

The results in this section indicate that, *ceteris paribus*, the choices of lesser educated individuals are associated with a larger noise parameter – since the preference parameters are constant across subjects, a larger noise parameter implies greater randomness in the decision-making process. Further, since access to formal education is indicative of the exposure to cognitive exercises and of cognitive ability (Ceci 1991), the results imply that the noise parameter varies systematically with cognitive ability, and are suggestive of the presence of poor understanding or confusion on the part of subjects (due to relative complexity of the MPL procedure). Humphrey and Verschoor (2004a) find that greater noise in decision-making is indicative of poor understanding, while Charness et al. (2013) also note that subjects with low education and numeracy skills, in particular, find the MPL procedure difficult to understand due to its greater complexity. Additionally, Barr and Packard (2005), analyzing the choices of participants in an experiment conducted in Peru, find evidence of limited comprehension of MPL problems involving insurance purchase decisions, which leads to greater observed noise in decision-making.

Further, for each decision model, the coefficient estimate of education is much larger than all other coefficient estimates except the estimate of the constant term (β_0), highlighting the importance of limited cognitive ability as a source of noisy behaviour in experimental decision-making (Dave et al. 2010).

It is important to note that the analyses of Dave et al. (2010) and Crosetto and Filippin (2013) assume EUT-CRRA preferences; thus, the interpretation of their estimates could be problematic if EUT-CRRA does not accurately represent the decision-making process of participants. That is, large noise parameter estimates could then simply indicate that the true preferences of participants in their experiment are not appropriately described by the EUT-CRRA model – and may, for example, be better described by non-EUT models – rather than reflecting randomness in choice due to poor understanding. In such a case, the assumption of EUT-CRRA would lead to biased estimates of preference parameters which do not accurately reflect true risk attitudes (Charness et al. 2013). This is a crucial issue and once again highlights the importance of considering a wide range of decision models to ensure the robustness of the results, and to confirm that results (for example, in relation to the noise parameter and the extent of errors in decision-making) are not driven solely by the functional form assumptions.

Further, Dave et al. (2010) compare estimates of the noise parameter μ across contexts in which the preferences parameters differ, without conducting any normalization procedure. They conduct structural maximum likelihood estimations (similar to those reported in this section) involving preference and noise parameters but allow for heterogeneity in both the preference and noise parameters across subjects – that is, they model both the preference and noise parameters as linear functions of individual characteristics. They find that the estimated preference parameters vary substantially by subject characteristics – however, they directly compare the estimates of μ obtained for subjects with different characteristics in order to determine which subjects make greater errors in decision-making; they interpret a larger magnitude of the estimated μ for a particular subject as representative of greater noise in the experimental decisions of that individual.[58] As highlighted earlier, comparing μ across contexts in which the preference parameters differ could lead to incorrect results and conclusions. Additionally, both Dave et al. (2010) and Crosetto and Filippin (2013) conduct different elicitation procedures with the same sample, and compare the extent of errors in decision-making across the different procedures through a comparison of the noise parameter estimates (using the Luce and Fechner error models), even though they find that the preference parameter estimates differ significantly across the different elicitation procedures.

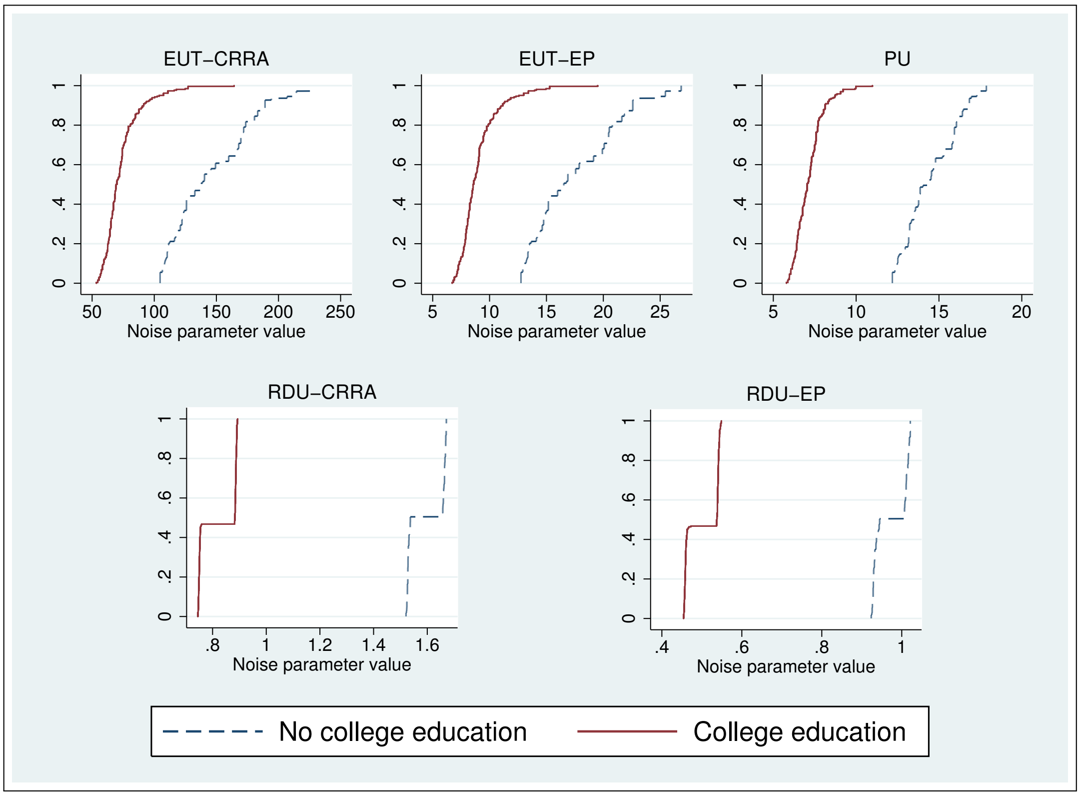
I present in Figure 7 graphs of the cumulative distribution of the predicted noise parameter in the Fechner error specification by college education (based on the estimations in this section), for each of the five decision models estimated.[59] This enables the visualization of the differences in the estimated noise parameter across college and non-college educted subjects. For the PU, RDU-CRRA and RDU-EP specifications, the entire range of the predicted noise parameter for participants who had completed – or were currently attending – college at the time of the experiment is lower than that for participants who did not attend college – there is no overlap in the distributions. Additionally, for the two EUT specifications, the extent of this overlap is very small.[60] This provides further evidence of the striking difference in the stochastic errors for the two groups, and further indicates that subjects with greater formal education – who are expected

[58] For completeness, I attempted to estimated similar specifications with both the noise and preference parameters defined as linear functions of the four individual characteristics considered in this section; however, the optimization routine did not converge for any of the decision models, likely due to the larger number of parameters to be estimated. Further, the analysis of Dave et al. (2010) focuses primarily on the EUT-CRRA model, which includes just one preference parameter, and they also encounter non-convergence when attempting to jointly estimate the correlates of the preference and noise parameters in expo-power and power utility specifications. Further, Crosetto and Filippin (2013) encounter convergence issues when attempting to jointly estimate the correlates of the risk aversion coefficient and the noise parameter in an EUT-CRRA model – as a result, they only estimate minimal specifications in which the noise parameter is a constant and the risk aversion coefficient is expressed as a linear function of gender. They hypothesize that these estimation issues arise because the choices in their decision problems cover a relatively small range of possibilities, as is common for experimental decision problems.

[59] Corresponding graphs for the Luce error specification are very similar.

[60] Dave et al. (2010), in similar graphs using data from MPL decision problems, also find virtually no overlap in the distributions of the predicted noise parameter for participants with higher cognitive ability and those with lower cognitive ability.

Figure 7. *Cumulative distribution of the noise parameter by college education*



to have higher cognitive ability (Ceci 1991) – exhibit considerably lower noise in the MPL decision problems.

The graphical representations – provided in Figures 1 and 2 – of the noise parameter estimates obtained in Section 4.1 indicate large deviations in observed choices from the predictions of the standard decision models considered, highlighting the possibility that none of these standard decision models accurately describe the true latent decision-making process of participants. To further investigate this possibility, I analyze "inconsistent" choices made by participants in the decision problems.

In the MPL decision problems of the Brazilian experiment, if a subject chooses No insurance in a particular question and then switches to Yes insurance in a later question (that is, a question lower in the table), this pattern is not consistent with concave or even convex preferences (Jacobson and Petrie 2007). Thus, any pattern of choices in the 20 questions of a particular decision problem in which the subject switches from choosing No to Yes insurance is “inconsistent”, following the nomenclature used in other experimental studies analyzing MPL data (for example, Harrison et al. 2005, Jacobson and Petrie 2009, Dave et al. 2010); such patterns are not consistent with any of the standard decision theories, or indeed with any well-defined preferences over risk. These inconsistent patterns involve indicating preference for not buying the insurance contract at a lower price but switching to indicating preference for buying the same insurance contract at a higher price (when the option of buying the insurance is unambiguously less attractive). These patterns involve dominated choices, non-monotonicity of revealed preferences and intransitivity in choice; on the other hand, all models of decision-making under risk with well-defined preferences over risk include transitivity as an intrinsic property (Starmer 1999). Thus, such inconsistent behaviour cannot be rationalized under standard assumptions on risk preferences (Charness et al. 2013).

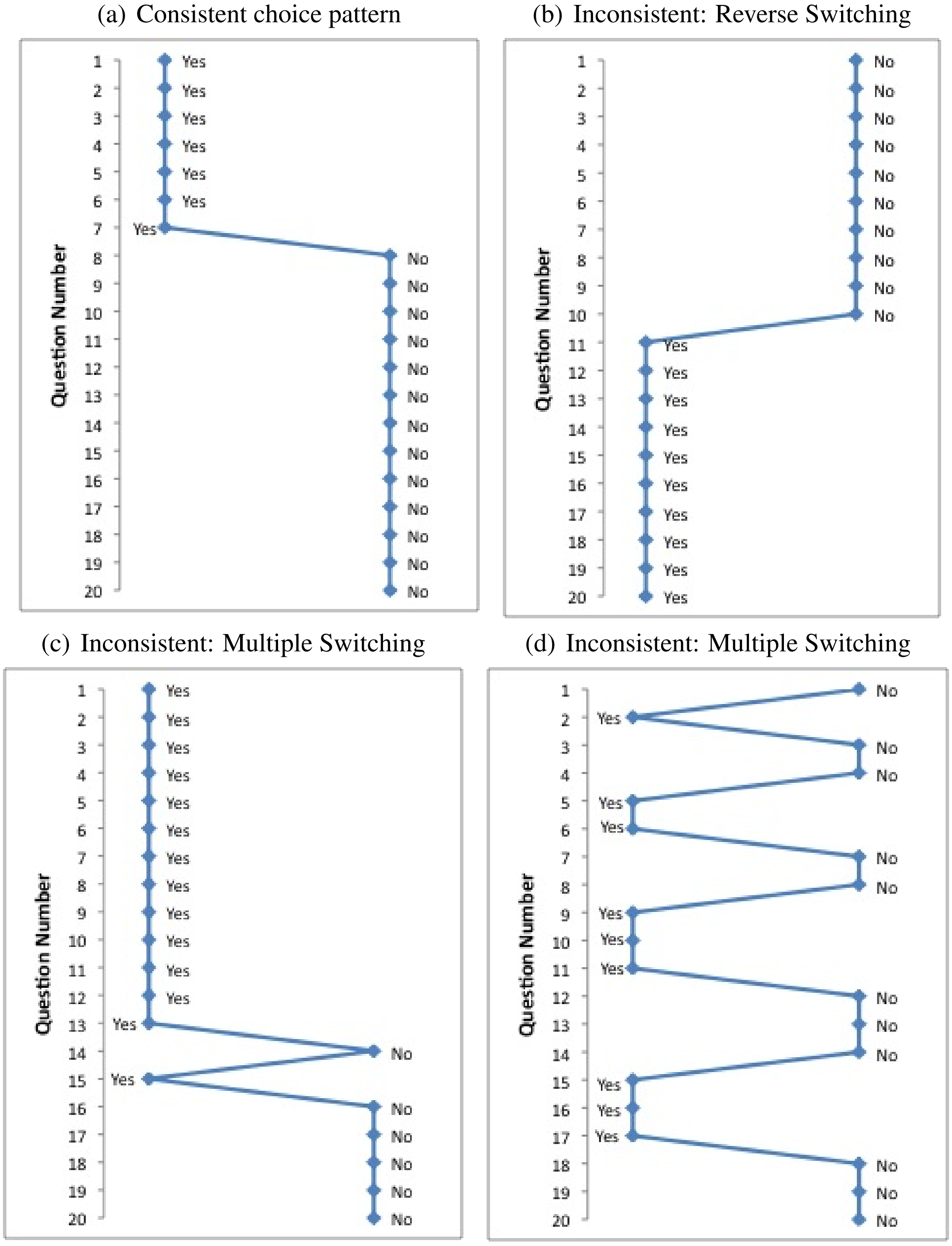
Consistent patterns include those in which participants choose the Yes insurance alternative at lower prices – since at lower prices the Yes insurance option has a higher net expected payoff – and then switch (once) to the No insurance option later in the table, that is, when the insurance contract is more expensive. The question number at which participants make this single switch also indicates their level of risk aversion (assuming EUT-CRRA), with more risk averse subjects switching from the Yes insurance option to the No insurance alternative lower in the table (at a higher price), and less risk averse subjects switching higher in the table (at a lower price). Barr and Packard (2005) note that in the case of consistent choices, the price at which a participant switched from buying to not buying insurance would reveal his preferences over risk. Risk neutral participants, for example, would be expected to switch when the expected values of the Yes and No insurance options are equal. Consistent (and transitive) patterns also include those in which the Yes option is chosen in all questions and those in which the No option is chosen in all questions.[61] Inconsistent patterns, on the other hand, include those beginning with preference for the No insurance alternative and then switching (once) to Yes insurance for the remainder of the questions ("reverse switching"), as well as those that involve switching between Yes and No insurance more than once ("multiple switching"). Figure 8 provides examples of some different types of choice patterns observed in the data – Panel (a) shows a consistent choice pattern and Panel (b) shows an inconsistent choice pattern with a single reverse switch, while Panels (c) and (d) provide examples of inconsistent choice patterns involving multiple switching.

In the seven MPL decision problems of the Brazilian experiment, 124 (or 33%) of the 374 choice patterns are inconsistent. Of these 124 inconsistent choice patterns, 122 involve multiple switching while only 2 involve a single reverse switch from No to Yes insurance. While such inconsistent, or intransitive, patterns are observed in nearly all MPL experiments, this percentage is higher than that obtained in many experiments conducted in developed countries – for example, Harrison et al. (2002) (3%, Denmark), Andersen et al. (2006b) (5.8%, Denmark and the United States), Dave et al. (2010) (8.5%, Canada), Meier and Sprenger (2007) (12%, United States) and Holt and Laury (2002) (13%, United States).[62] However, the fraction of inconsistent choice patterns observed in the Brazilian experiment, while relatively high, is still in line with that found in

[61] Strictly speaking, the choice pattern in which the Yes insurance option is chosen in every question is consistent as it does not involve any intransitive choices, and in this chapter I equate consistency with transitivity. However, in question 20, the Yes insurance option is dominated by the No insurance option – with the insurance contract priced at R$8, there is no net gain from the contract when Box A is chosen, but there is still a net loss (of R$8) from the contract when Box B is selected. Thus, if the decision problems in the experiment were properly understood (and in the absence of stochastic errors), every participant should have chosen the No insurance option in question 20; as a result, the presence of a large number of choice patterns in which the Yes insurance option is chosen in every question (including question 20) could be indicative of poor understanding (Charness and Viceisza 2012). In the analyses conducted in this chapter, I define the pattern of all Yes choices (observed 19 times) as consistent (as done by Harrison et al. 2007), since these choices are strictly transitive. However, even if this pattern is classified as "inconsistent", as done by Charness and Viceisza (2012), there are no significant differences in the results obtained in this and the following section.

[62] It is important to note that while such inconsistent patterns (due to multiple or reverse switching) are commonly observed in data from MPL decision problems, there are some variations of the MPL design which impose consistency in choices (Andersen et al. 2006b); however, like the original Holt and Laury (2002) elicitation procedure commonly used in experiments, the procedure used in the Brazilian experiment does not impose any such restrictions on subjects' choices.

Figure 8. *Examples of observed consistent and inconsistent choice patterns*



other MPL experiments, particularly those conducted in developing countries – Jacobson and Petrie (2009), for example, find 55% and 50% of choice patterns to be inconsistent in MPL experiments conducted in Rwanda and Peru, respectively; around half of the choice patterns in Charness and Viceisza's (2012) experimental dataset from rural Senegal are intransitive, while Doerr et al. (2011) observe an inconsistency rate of 39% in Ethiopia. Additionally, Laury and Holt (2005) and Prasad and Salmon (2007) find that almost 30% of their samples of participants from the United States make inconsistent choices in MPL decision problems.[63]

The significant fraction of inconsistent choice patterns observed in the Brazilian experiment could be a cause of the substantial randomness in choice estimated in the Luce and Fechner error models in previous section, as these patterns are not in line with the predictions of the standard decision models considered (and indeed with any well-defined preferences over risk). Additionally, Charness et al. (2013) and Crosetto and Filippin (2013) note that a large fraction of inconsistent choice patterns is indicative of poor understanding of the complex MPL decision problems. However, Barr (2007) lists two other possible causes of the inconsistent choice patterns – random errors in decision-making and genuine indifference between the alternatives in a decision problem – in addition to poor understanding. An important aim of this chapter is to determine which of these three factors is the most likely cause of the inconsistent patterns observed in the

[63] It is important to bear in mind that in an Ordered Lottery Selection problem, subjects make only a single choice, and thus this elicitation procedure does not allow inconsistent choices to be made (in a single decision problem), by design.

Brazilian experiment.[64]

While both types of multiple switching patterns illustrated in Figure 8 (Panels (c) and (d)) are intransitive, they may reflect very different underlying decision-making behaviour. Patterns in which subjects switch only twice from the Yes to the No option (as displayed in Panel (c)) – particularly over a short span of questions – could indicate that participants exhibiting such patterns are indifferent between the options within the range of questions over which these switches occur. Andersen et al. (2006b), analyzing experimental data from Denmark and the United States, conclude that indifference is the most likely cause of the multiple switching behaviour often observed in MPL experiments (due to the non-inclusion of an explicit option to indicate indifference). In such a case, the choice patterns could still yield valuable information about risk preferences, and wider ("fatter") intervals – determined by the first and last rows at which the subject switched – of preference parameters could be used to represent subjects' risk attitudes. This strategy has been commonly used in the experimental

[64] Another possible cause of intransitivity in experimental choice, as noted by Tversky and Kahneman (1974), is the use of the anchoring and adjustment decision heuristic combined with insufficient (or imperfect) adjustment. However, it is unlikely that the intransitivity in choice observed in the Brazilian experiment stems from this, since the heuristic is generally considered to explain decision-making under uncertainty, rather than decision-making under risk. For example, Epley and Gilovich (2006) note that "one way to make judgments under uncertainty is to anchor on information that comes to mind and adjust until a plausible estimate is reached". Further, Humphrey (1996) tests the presence of anchoring and adjustment using an experiment in which complete information on the probabilities associated with different outcomes is not given to participants – in such a context, it is possible that in "forming probability assessments individuals begin by forming an anchor assessment on the basis of, for example, the formulation of the problem and then adjust this anchor to reach a final probability assessment". However, in the Brazilian experiment, participants were given complete information about the outcomes and associated probabilities in the decision problems. In addition, as noted by Humphrey (1996), anchoring and insufficient adjustment can lead to intransitive choices, but the intransitive choices only occur in a particular direction. For example, in his experiment, the direction of violation of transitivity due to anchoring and insufficient adjustment is such that participants are expected to exhibit less risk aversion when complete information on the probabilities is not given as compared to when it is (that is, "less risk aversion under uncertainty than under risk"). However, in the Brazilian experiment, the observed intransitive choices are not in any one particular direction (that is, towards more or less risky options). Rather, as noted later, most of the observed intransitive choice patterns involve repeated switching over a large range of questions; that is, intransitive choices occur in both directions, indicating that subjects may not have properly understood the decision problems and instead selected choices in a psuedo-random, arbitrary fashion – for example, 80 of the 124 inconsistent choice patterns involve five or more switches between the Yes and No insurance options. Thus, the switching does not occur in a predictable manner or (predominantly) in a single direction (which could be explained by anchoring and imperfect adjustment).

literature to characterize subjects with such choice patterns (for example, Coller et al. 2003, Harrison et al. 2003). However, the experimental datasets analyzed in these studies contain only a small fraction of observed intransitive choice patterns; further, of these intransitive patterns, most involve only two switches from the safe to the risky option.

On the other hand, in the Brazilian experiment, only 18 of the 122 observed multiple switching patterns involved just two switches from the Yes to No insurance option. The remaining multiple switching patterns involved more than two switches from Yes to No insurance, and many of these included a large number of switches over a wide range of questions, as shown in the pattern in Panel (d) of Figure 8. Such patterns involving repeated switching over a wide range of questions are likely to reflect a fundamental lack of understanding of the decision problems, as opposed to indifference between alternatives (Coller et al. 2003). Thus, the relatively large number of patterns observed in the data which involve repeated switching between the Yes and No alternatives over a wide range of questions indicates poor understanding of the decision problems in the Brazilian experiment on the part of subjects, rather than indifference between alternatives.[65,66]

Building on the analysis of the correlates of the noise parameter earlier in this section, and in line with the work of Jacobson and Petrie (2007), I conduct Probit

[65] Further, since 122 of the 124 intransitive choice patterns observed in the data involve multiple switching, the indication that the multiple switching patterns are unlikely to be caused by indifference also implies that the inconsistent choice patterns are unlikely to result from indifference between alternatives.

[66] Some studies consider the last "safe" choice of subjects to represent true preferences over risk, and use this measure in analyses and estimations of risk preferences (Andersen et al. 2006b). However, using the last Yes choice to represent true risk preferences or the willingness to pay for insurance in the experimental data at hand would be inappropriate. This is because a large fraction of observed choice patterns are intransitive, and most of these intransitive patterns involve repeated switching over a large range of questions, indicating that subjects may not have properly understood the decision problems and instead selected choices in a psuedo-random fashion – for example, 80 of the 124 inconsistent choice patterns involve five or more switches between the Yes and No insurance options. Thus, the last question at which the subject chose the Yes insurance option may not provide meaningful information about risk preferences, but could instead be a product of randomness in choice. Similarly, another commonly used method to summarize a MPL decision problem, the total number of "safe" choices, is not likely to provide meaningful information in this case (a point which is also made by Dave et al. 2010).

regressions of a binary dummy indicating whether the choice pattern in a decision problem was inconsistent or not on the four individual characteristics – age, gender, education and wealth. This enables the evaluation of the relationship between inconsistencies in choice and these characteristics.[67] Additionally, following Jacobson and Petrie (2007), I also conduct Ordered Probit regressions of the number of inconsistent choices on these four variables, where the number of inconsistencies (or inconsistent choices) for a particular choice pattern in a decision problem is defined as the number of reverse switches from the No insurance to the Yes insurance option.[68] The results of these regressions are presented in Table 12.[69,70]

Column (1) of Table 12 presents the coefficient estimates for the Probit regression, while Column (2) presents the corresponding marginal effects for this specification, evaluated at the means of the four explanatory variables. The results in these columns indicate that participants who did not attend college were more likely to have an inconsistent choice pattern, in line with the results in Tables 10 and 11, which show that lesser education is associated with a larger noise parameter estimate in the Luce and Fechner error models. The coefficient estimate of college education in Column (1) is negative and

[67] An inconsistent choice pattern is defined as one in which there is at least one switch from the No insurance to the Yes insurance option. The binary inconsistent dummy therefore equals one when there is at least one such reverse switch present in a particular choice pattern.

[68] Thus, the number of inconsistencies is zero for consistent choice patterns, and the maximum number of inconsistencies for any choice pattern is ten.

[69] I also conducted Logit and Ordered Logit regressions, which yield very similar results to the corresponding Probit and Ordered Probit regressions. All these models also utilize maximum likelihood estimators.

[70] Most studies drop patterns involving any form of reverse or multiple switching from estimations (Coller et al. 2003). However, following the strategy of Jacobson and Petrie (2009), I do not drop any of the observed intransitive choice patterns, as this could lead to a loss of important information on the decision-making under risk of participants. Additionally, Charness and Viceisza (2012) note that if a large fraction of subjects make inconsistent choices – as in the case in the Brazilian experiment – then dropping the inconsistent subjects or choices from the analysis is not a viable option and could lead to skewed results (that are also inappropriate for making policy recommendations). Further, if inconsistent choices are related to lower cognitive ability (as the results of this analysis indicate), then removing these observations from the data would significantly affect the characteristics of the sample and bias the estimates of risk preferences (Charness et al. 2013). However, I cannot discount the possibility that subjects who did not understand the decision problems still indicated a transitive choice pattern by chance.

Table 12. *Probit and Ordered Probit regressions of inconsistency measures*

Variables	Probit (1)	Probit (Marginal Effects) (2)	Ordered Probit (3)
Age	0.0242**	0.00861**	0.0304*
	(0.0123)	(0.00440)	(0.0180)
Gender	-0.0318	-0.0113	0.0187
	(0.141)	(0.0503)	(0.130)
College education	-0.809***	-0.301***	-0.766***
	(0.171)	(0.0634)	(0.157)
PCA wealth	0.237**	0.0845**	0.215*
	(0.121)	(0.0431)	(0.112)
Constant	-0.716*		15.90***
	(0.396)		(5.573)
Observations	372		372
Log-likelihood	-216.4		-441.1

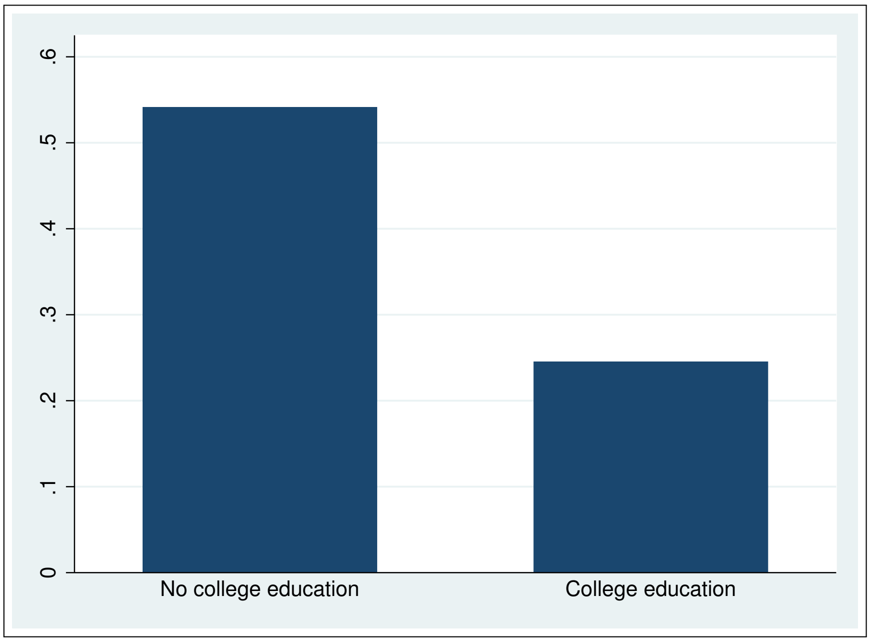
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$
Dependent variable in the Probit regression is a binary dummy which equals one if the choice pattern is inconsistent and zero if it is consistent. Dependent variable in the Ordered Probit regression is the number of inconsistencies in the choice pattern (between 0 and 10). Standard errors are in parentheses.

statistically significant at the 1% level.[71] The marginal effect of education on the probability of having an inconsistent choice pattern (Column (2)) is also negative and statistically significant at the 1% level; further, the magnitude of this marginal effect is estimated to be relatively large – the estimate indicates that participants who did not attend college had a 30% greater chance of making at least one reverse switch than those who did. Dave et al. (2010) also find that a significantly larger fraction of low math ability subjects in their experiment make inconsistent choices in a MPL problem as compared to high math ability subjects.

The Probit results are visualized in Figure 9, which separately plots the fractions of inconsistent choice patterns made by college and non-college educated subjects – the figure shows that there is a considerably higher fraction of inconsistent choice patterns among participants who did not attend college (54%) as compared to the fraction of inconsistent choice patterns among college-educated subjects (25%). Further, in the Ordered Probit regression (Column (3)), the coefficient

[71] Jacobson and Petrie (2007) also find a negative association between education and the probability of making an inconsistent choice, in a similar Probit regression.

Figure 9. *Fraction of inconsistent choice patterns by college education*



estimate of college education is once again negative and statistically significant (at the 1% level), indicating that participants who did not attend college made a larger number of inconsistent choices in the decision problems, *ceteris paribus*.[72]

The results in Table 12 also indicate that older participants were more likely to have an inconsistent choice pattern (Probit regression) as well as a greater number of inconsistencies (Ordered Probit regression). However, the estimated magnitude of the marginal effect of age on the probability of having an inconsistent choice pattern is relatively small. While the association between wealth and the noise parameter was not statistically significant in the Luce and Fechner error specifications (Tables 10 and 11), wealth has a positive and statistically significant coefficient estimate in both the Probit and Ordered Probit regressions. Further, the results in Column (2) indicate that a one unit increase in the PCA wealth index is associated with an 8% increase in the probability of having an inconsistent choice pattern.

The finding that education is negatively related to both the noise parameter and inconsistencies could indicate that participants with lower cognitive ability faced

[72] In the Ordered Probit regression, the estimated thresholds (cutoffs) 1 through 10 are 0.903, 1.205, 1.452, 1.634, 1.904, 2.385, 2.658, 2.783, 3.060 and 3.493.

difficulty in understanding the decision problems, and thus were more likely to make inconsistent choices, as well as have greater stochastic errors in their decision-making (Dave et al. 2010, Crosetto and Filippin 2013).[73] It is likely that greater access to education increases exposure to cognitive exercises, which may lead to better comprehension of the experimental decision problems. Access to formal education is generally more abundant in developed than developing economies (Jacoby and Skoufias 1997), and therefore individuals with low formal education are likely to comprise a significant fraction of experimental samples in underdeveloped regions of developing countries – as a result, simpler experimental procedures, rather than the relatively complex MPL mechanism, may be better understood and more appropriate for accurately eliciting risk preferences in experiments conducted in these regions. Additionally, simpler procedures may also be more suitable for eliciting the risk preferences of non-standard experimental samples – consisting of individuals who are not university students – in developed countries; these samples may also be characterized by relatively low mathematical skills (Dave et al. 2010, Charness et al. 2013).[74]

However, it is important to consider the possibility that the inconsistent choices do not arise due to poor understanding, but rather due to random errors in decision-making as a result of, for example, miscalculations or lapses in concentration, which cause observed choices to deviate from the true (transitive) risk preferences. This possibility is explored in the following section.

4.3 Alternative decision models and constant error rate analysis

The results of the previous section raise the following question: are intransitive choices, and the large estimated extent of randomness in choice, a result of

[73] The finding in the Ordered Probit regression that subjects with lower cognitive ability made a larger number of reverse switches indicates that these subjects were more likely to have patterns involving repeated switching between the Yes and No insurance option. This highlights the possibility that the large number of repeated switching patterns observed in the data are caused by poor understanding rather than indifference between alternatives.

[74] On the other hand, Dave et al. (2010) find that for experimental samples consisting of individuals with relatively high levels of cognitive or numerical ability, the finer characterization of risk preferences enabled by the complex MPL procedure outweighs the greater noise.

mistakes (or errors) which cause deviations in observed choices from the true decision-making process which is transitive (and possibly given by one of the standard models), or is the latent decision-making process of participants non-transitive and better described by alternative decision models which generate intransitive choice patterns? To address this question, I first use binomial tests to evaluate whether subjects make choices by randomizing with equal probability between the two options in a question, and then use the constant error rate analysis of Harless and Camerer (1994) to test whether observed intransitive choice patterns are a result of the true decision-making process or random mistakes. Finally, I estimate, and consider the suitability in explaining the observed choices of, two alternative decision models. These models are markedly different from the standard decision models analyzed so far in their setup and assumptions, and do not conform to any well-defined preferences over risk – thus, the use of such decision-making processes is likely due to poor comprehension of the decision problems.

If all participants truly did not understand – and were completely confused by – the decision problems, they would be expected to make choices entirely at random. One likely randomization procedure in this case is to randomize with equal probability (of 0.5) between the two options in a question. To explore this possibility, I formally test the hypothesis that the observed choices were generated by this randomization procedure. In order to do this, I combine data from all seven decision problems and use binomial tests to compare the observed fractions of Yes insurance choices in each of the 20 questions to 0.5. The null hypothesis is that the observed fractions of Yes choices in all 20 questions are 0.5, which would be the case if choices in the questions were made completely at random with equal probability of choosing each alternative. Thus, if the binomial test rejects that the observed fraction of Yes choices is 0.5 for even one of the questions, it implies a rejection of the hypothesis that the binary choices were generated at random with equal probability for each alternative. Using this test, I find that this hypothesis can be rejected (at a significance level of 10% or lower) for all questions except questions 7 and 8. This indicates a rejection of the hypothesis of randomization with equal probability. However, this does not provide conclusive evidence against poor comprehension, as participants could have used other heuristics or randomization strategies, which are also indicative of poor understanding, to make decisions. This possibility is explored in the following analyses.

To test the validity of transitive choice models (which include all the deterministic standard decision models considered earlier), as a whole, for explaining the observed choices in the Brazilian experiment, I utilize the constant error rate analysis proposed by Harless and Camerer (1994). As done by Sopher and Gigliotti (1993), I use this test to determine whether the intransitive choice patterns observed in the data are in accordance with the true decision-making process of participants, or due to random errors (referred to as "trembling hand" errors) which cause deviations in observed choices from true preferences, which are transitive. Similar tests involving Harless and Camerer's (1994) constant error model have been used by Humphrey and Verschoor (2004a) to study preference reversals in experimental decision problems conducted with farmers in East Uganda, and by Jacobson and Petrie (2009) to analyze errors in decision-making in an experiment conducted in Rwanda. Further, this constant error rate analysis is non-parametric, enabling abstraction from functional form issues. Additionally, numerous studies have used the error structure proposed by Harless and Camerer (1994) – combined with a maximum likelihood estimation procedure – to model and evaluate individual behaviour in experimental decision problems (for example, Loomes et al. 2002).

There are 20 binary questions in each decision problem of the Brazilian experiment, which implies that there are 2^{20} (=1,048,576) possible types of choice patterns. However, only 21 of these choice patterns are transitive (and thus consistent with deterministic models of decision-making under risk involving well-defined risk preferences) – these include the patterns which involve all Yes choices, all No choices, and the 19 patterns which involve a single switch from the Yes to No insurance choice.[75] All the remaining choice patterns involve intransitive choices.

In this constant error rate analysis, it is first assumed that each subject's true preferences are transitive – and thus given by one of the 21 transitive choice patterns – but in any given question of a decision problem there is a fixed (constant) probability (ϵ) with which subjects make a mistake and do not state their true preferences (Sopher and Gigliotti 1993). Thus, all subjects have true preferences that are transitive, and intransitive choice patterns are only generated due to mistakes.[76] Following Harless and Camerer (1994), the probability of error (or error rate) ϵ is assumed to be constant across subjects and across questions – that is, it is assumed to independent and equal.[77,78] It is important to note that in this setup, all intransitive patterns are not, in general, equally likely, and the distribution of intransitive patterns depends on the distribution of transitive patterns (Sopher and Gigliotti 1993); for example, if there are a large number of participants whose true preferences are given by the transitive pattern in which

[75] The 19 single switch patterns consist of those in which subjects switch from Yes to No insurance at question 2 (that is, choose Yes in question 1 and then No in questions 2-20), those in which subjects switch at question 3 (that is, choose Yes in questions 1-2 and then No in questions 3-20), and so on till the pattern in which subjects switch at question 20.

[76] It is important to note that the preferences of subjects can be given by any of the 21 transitive choice patterns, and different subjects can have true preferences corresponding to different transitive choice patterns. Therefore, in this analysis, the representative agent assumption – which implies that all individuals have the same preferences – is dropped. Further, in this analysis, parametric forms are not imposed on the preferences of subjects, and I investigate the validity of transitive preferences as a whole (which encompass all the deterministic standard decision models considered earlier), rather than the suitability of a particular form of preference function. Thus, using the constant error rate analysis, a rejection of transitive preferences would indicate a rejection of all the standard decision models considered, and indeed of any model of decision-making under risk (with well-defined preferences over risk).

[77] Random errors of this form are referred to in the literature as "trembling hand" errors, which could arise due carelessness, miscalculations or a lapse of concentration on the part of subjects (Loomes 2005, Humphrey and Verschoor 2004a). The "trembling hand" error model of Harless and Camerer (1994) is different from the "white noise" error model of Hey and Orme (1994) (which was used in the construction of the Luce and Fechner specifications). Both models allow for deviations in observed choices from true preferences and have been used extensively to model errors in experimental decision-making (Loomes et al. 2002). However, unlike the "white noise" error model, the "trembling hand" model involves a constant probability of error in each decision and also does not require parametric or functional form assumptions.

[78] Harless and Camerer (1994) note that using a single constant error rate provides an appropriate parsimonious approach for explaining the distribution of choice responses, and find no form of error dependence persuasive. Further, they find that allowing error rates to be choice dependent in such a model can produce nonsensical results.

the unique switch from Yes to No insurance occurs at question 5, then one would expect to observe more intransitive patterns which involve a single choice deviation (that is, an error in a single question) from this transitive pattern than those which involve two or more choice deviations (that is, errors in two or more questions).

Using the assumption that the true preferences of subjects are given by the transitive choice patterns, and intransitive patterns are observed only due to mistakes, the first part of this test involves using maximum likelihood to estimate the error probability ϵ as well as the proportion of individuals in the sample with true preferences corresponding to the different transitive choice patterns. Of the 21 transitive choice patterns, two patterns are only observed once in the data – switching (once) at question 15 and switching (once) at question 19. Harless and Camerer (1994) note that the maximum likelihood estimation procedure could encounter convergence issues when attempting to estimate the proportions associated with patterns that are observed very infrequently (or not observed) in the data. Further, Sopher and Gigliotti (1993) also estimate only the most prominent share parameters, and omit transitive patterns with the lowest observed frequencies from their estimations, for parsimony and identification reasons. Thus, I also drop from my estimations the two patterns which occur only once in the data, and only consider the remaining 19 transitive patterns as representative of the true preferences of individuals. In other words, the true preferences of each participant are assumed to be given by one of these 19 transitive patterns, and any other observed patterns are generated by errors that cause deviations in observed choices from these true preferences.

Following the terminology used by Sopher and Gigliotti (1993), the probability that an individual i whose true preferences are given by pattern j *appears* to be of type m (where m can be a transitive or intransitive pattern) is given by:

$$
\begin{aligned}
\pi_{ij}^{m} &= (1-\epsilon)^{D_{ij1}}\epsilon^{1-D_{ij1}}(1-\epsilon)^{D_{ij2}}\epsilon^{1-D_{ij2}}(1-\epsilon)^{D_{ij3}}\epsilon^{1-D_{ij3}}(1-\epsilon)^{D_{ij4}}\epsilon^{1-D_{ij4}} \\
&(1-\epsilon)^{D_{ij5}}\epsilon^{1-D_{ij5}}(1-\epsilon)^{D_{ij6}}\epsilon^{1-D_{ij6}}(1-\epsilon)^{D_{ij7}}\epsilon^{1-D_{ij7}}(1-\epsilon)^{D_{ij8}}\epsilon^{1-D_{ij8}} \\
&(1-\epsilon)^{D_{ij9}}\epsilon^{1-D_{ij9}}(1-\epsilon)^{D_{ij10}}\epsilon^{1-D_{ij10}}(1-\epsilon)^{D_{ij11}}\epsilon^{1-D_{ij11}}(1-\epsilon)^{D_{ij12}}\epsilon^{1-D_{ij12}} \\
&(1-\epsilon)^{D_{ij13}}\epsilon^{1-D_{ij13}}(1-\epsilon)^{D_{ij14}}\epsilon^{1-D_{ij14}}(1-\epsilon)^{D_{ij15}}\epsilon^{1-D_{ij15}}(1-\epsilon)^{D_{ij16}}\epsilon^{1-D_{ij16}} \\
&(1-\epsilon)^{D_{ij17}}\epsilon^{1-D_{ij17}}(1-\epsilon)^{D_{ij18}}\epsilon^{1-D_{ij18}}(1-\epsilon)^{D_{ij19}}\epsilon^{1-D_{ij19}}(1-\epsilon)^{D_{ij20}}\epsilon^{1-D_{ij20}}
\end{aligned}
\tag{20}
$$

Thus,

$$\pi_{ij}^{m} = (1-\epsilon)^{\sum_{k=1}^{20} D_{ijk}}(\epsilon)^{20-\sum_{k=1}^{20} D_{ijk}} \tag{21}$$

where D_{ijk} is equal to zero if an individual i of type j would need to make a mistake in answering question $k \in \{1, \ldots, 20\}$ in a decision problem to appear of type m, and is equal to one otherwise (Sopher and Gigliotti 1993). Thus, the probability of observing a pattern m (transitive or intransitive) for individual i – not conditioning on the true type of the individual – is given by:

$$
\begin{aligned}
\pi_{i}^{m} &= P_2\pi_{i2}^{m} + P_3\pi_{i3}^{m} + P_4\pi_{i4}^{m} + P_5\pi_{i5}^{m} + P_6\pi_{i6}^{m} + P_7\pi_{i7}^{m} + P_8\pi_{i8}^{m} \\
&+ P_9\pi_{i9}^{m} + P_{10}\pi_{i10}^{m} + P_{11}\pi_{i11}^{m} + P_{12}\pi_{i2}^{m} + P_{13}\pi_{i13}^{m} + P_{14}\pi_{i4}^{m} \\
&+ P_{16}\pi_{i16}^{m} + P_{17}\pi_{i17}^{m} + P_{18}\pi_{i18}^{m} + P_{20}\pi_{i20}^{m} + P_{21}\pi_{i21}^{m} + P_{22}\pi_{i22}^{m}
\end{aligned}
\tag{22}
$$

where P_j is the proportion of subjects in the sample whose true type is j. For $j \in \{2, \ldots, 20\}$, P_j indicates the proportion of participants who have true preference patterns in which they switch once at question j (that is, choose Yes insurance up to question $j-1$ and then No insurance from question j onwards); P_{21} and P_{22} refer to the proportions of participants whose true preference patterns involve all No choices and all Yes choices, respectively.[79] Further, following Sopher and Gigliotti (1993), I impose the restriction $P_{22} = 1 - \sum_{j=2}^{14} P_j - P_{16} - P_{17} - P_{18} - P_{20} - P_{21}$. The log-likelihood function is thus the sum of the natural

[79] Note that P_{15} and P_{19} are omitted from this specification, since the corresponding patterns are only observed once (each) in the data.

log of π_i^m over all individuals and all decision problems. This log-likelihood function can then be maximized with respect to the 19 parameters to be estimated – that is, 18 proportion (P) parameters and the error parameter (ϵ) – to obtain the maximum likelihood estimates of these parameters. The results of this estimation are presented in Column (1) of Table 16 in Appendix A.[80]

The second part of this test involves allowing intransitive patterns to represent "true" preference patterns as well. The decision problems considered by Sopher and Gigliotti (1993) and Harless and Camerer (1994) consisted of only three binary choices – thus, there were only eight possible choice patterns, six of which were transitive. As a result, they only had to consider two intransitive choice patterns, and their estimations assumed these two patterns to be generated by mistakes causing deviations from the six transitive patterns. However, in each of the decision problems of the Brazilian experiment, there were 20 binary questions, and thus 1,048,555 (=1,048,576−21) possible intransitive patterns. Therefore, it is unfeasible to model and account for all the possible intransitive patterns as representative of true preferences. Instead, for estimation purposes, I only allow the five intransitive patterns which occur more than once in the experimental data to reflect true preferences. Thus, I now assume that in addition to the 19 transitive patterns, the five intransitive patterns (which occur more than once in the data) can also represent the true preferences of individuals.

The expression for π_i^m described in Equation (22) is modified to allow for this alternative hypothesis that the intransitive patterns are also possible true preference patterns (Sopher and Gigliotti 1993), by adding $P_{23}\pi_{i23}^m + P_{24}\pi_{i24}^m + P_{25}\pi_{i25}^m + P_{26}\pi_{i26}^m + P_{27}\pi_{i27}^m$ to the expression, where these five proportions represent the proportions associated with the five intransitive patterns. As before, the log-likelihood function is the sum of the natural log of this modified π_i^m over all individuals and all decision problems. This function can then be maximized

80 Following Sopher and Gigliotti (1993) and Harless and Camerer (1994), the population share parameters are restricted to be between 0 and 1, and ϵ is restricted to be between 0 and 0.5 in these estimations. These maximum likelihood estimations were carried out in MATLAB using the sequential quadratic programming (SQP) optimization algorithm.

with respect to the 23 proportion parameters and the error parameter (ϵ) to obtain maximum likelihood estimates of these parameters. The results of this estimation are reported in Column (2) of Table 16 in Appendix A.[81] In these estimations, the true preferences of each participant are assumed to be given by one of these 24 choice patterns, and any other observed patterns are generated by errors that cause deviations in observed choices from these true preferences.

Using these two maximum likelihood estimations, I then perform a likelihood ratio test (LRT) to evaluate the null hypothesis, which is the restriction $P_{23} = P_{24} = P_{25} = P_{26} = P_{27} = 0$. Thus, the null hypothesis is that none of the five intransitive choice patterns represent true preferences (and thus the fractions of participants having these choice patterns as their true preferences are 0); in other words, participants have true preferences which are transitive, and the five intransitive patterns are only observed due to "trembling hand" errors in decision-making. Using the two models estimated – the restricted (first) and unrestricted (second) – the LRT statistic is calculated as follows:

$$2[\ell_{UR} - \ell_R] \sim \chi^2(r) \tag{23}$$

where ℓ_{UR} and ℓ_R are the optimized values of the log-likelihood functions for the unrestricted and restricted models, respectively, and r is the number of independent restrictions imposed in the restricted model.

Using the log-likelihood values from the maximum likelihood estimations (reported in Table 16), the LRT statistic value obtained is 134; further, since there are five restrictions involved in the restricted model, the LRT statistic is distributed χ^2 with five degrees of freedom. Thus, under the null hypothesis, the p-value obtained is very close to 0; the null hypothesis that $P_{23} = P_{24} = P_{25} = P_{26} = P_{27} = 0$ is convincingly rejected, implying that at least one of the five intransitive choice patterns represents true preferences. In other words, I can

[81] Note that in this estimation, I impose the restriction $P_{22} = 1 - \sum_{j=2}^{14} P_j - P_{16} - P_{17} - P_{18} - P_{20} - P_{21} - \sum_{j=23}^{27} P_j$.

reject the hypothesis that all five intransitive patterns considered are generated solely by "trembling hand" errors and all participants have true preferences that are transitive. Further, the estimates of the intransitive pattern proportion parameters P_{24} and P_{25} are statistically significantly different from 0 at the 5% level, providing evidence that these two patterns, in particular, could represent true preferences.[82]

The estimates of the constant error parameter (or error rate) ϵ are around 0.07 in both models (restricted and unrestricted) – both these estimates are statistically significantly different from 0, indicating the presence of some "trembling hand" errors in the decision-making process of participants; the magnitude of these estimates is in line with that of the estimates obtained by Sopher and Gigliotti (1993) and Harless and Camerer (1994).[83] Harless and Camerer (1994) note that this non-parametric analysis provides an appropriate complement to a parametric analysis of the form conducted in Section 4.1, in which I analyze various preference functions and "white noise" error models.[84] Additionally, this non-parametric approach possesses a great deal of flexibility, as it allows for heterogeneity in preferences – since different subjects are allowed to have different true preference patterns – and does not specify any preference or utility functions.

Such an analysis of transitive and intransitive patterns has been widely used in experiments involving relatively few choices – for example, Harless and Camerer (1994) conduct similar analyses using data from a number of different experiments, the largest involving five choices (and thus 32 possible choice patterns) in each decision problem. However, to my knowledge, an analysis of

[82] Sopher and Gigliotti (1993), on the other hand, cannot reject the null hypothesis that the addition of the intransitive patterns does not improve the fit of the model to the data, and conclude that the intransitive choice patterns observed in their experiment are generated by subjects with transitive preferences who make mistakes.

[83] However, these estimates of the constant error probability are lower than those obtained by Humphrey and Verschoor (2004a).

[84] Barr and Packard (2005) note the importance of conducting non-parametric analyses, which do not require making assumptions about the form of participants' utility functions, when evaluating decision-making under risk.

this type has not been previously conducted using MPL data involving a relatively large number of choices in each decision problem, such as in the Brazilian experiment. It is important to note that in this case, since the test only allows for a very small fraction of possible intransitive choice patterns to represent true preferences, it is a relatively conservative test. While I cannot realistically allow for all the intransitive patterns to represent true preferences in the maximum likelihood estimation of the unrestricted model, if more intransitive patterns – in addition to the five considered – are allowed to represent the true preferences of subjects, it would only serve to strengthen the result obtained in favour of intransitive patterns representing true preferences. Since the null hypothesis is already convincingly rejected (with a p-value very close to 0) with just the five intransitive patterns considered, there is strong evidence that the intransitive patterns reflect the true preferences and decision-making process of subjects, and are not simply generated by random errors.[85]

Thus, using the constant error rate analysis, I find evidence that the true decision-making process of participants may be intransitive. This provides some indication that the true decision-making process of subjects may not be appropriately described by any model of rational behaviour under risk (with well-defined risk preferences), including the (deterministic) standard decision models considered earlier, all of which assume transitivity in choice. Additionally, poor understanding of the complex MPL problems is a likely cause of the use of a decision-making process which generates intransitive and dominated choices.

I next estimate two alternative decision models that do not assume transitivity and thus allow the significant proportion of intransitive (or inconsistent) patterns observed in the data to be generated by the true decision-making process (rather than random errors). I then test whether these models better fit the data and describe observed choices than the transitive standard decision models considered in Section 4.1.

[85] It is important to note that the order in which binary lotteries (of a decision problem) were presented did not vary across subjects.

The first of these alternative models assumes the true latent decision-making process of participants to be random – that is, subjects randomize between the two alternatives in each binary question of a decision problem (hereafter, I refer to this model as the "Random Choice" model). In these estimations, participants are implicitly assumed to make choices at random in accordance with their true decision-making process, and not because of errors causing deviations in observed choices from true preferences. The second of these alternative models assumes that participants, when responding to a Multiple Price List decision problem, select their choice (out of the two options) in the first binary question with a particular (fixed) probability; then, in subsequent binary questions, they stick to their previous choice with some other (fixed) probability and with some probability switch their choice to the other option (hereafter, I refer to this model as the "Random Switching" model).

The Random Choice model involves only a single parameter to be estimated – the probability of choosing the Yes insurance option in a question (p_Y). To estimate this probability, I use data for all participants and all questions in all seven decision problems, and use a maximum likelihood estimation procedure in which the likelihood function is simply equal to p_Y if the Yes insurance option was chosen in a question, and $1 - p_Y$ if the No insurance option was chosen. The results of this estimation are presented in Panel A of Table 13. The Random Switching model, on the other hand, involves two parameters to be estimated – the probability of choosing the Yes insurance option in question 1 (p_{Y1}), and in subsequent questions (2 to 20), the probability of choosing the same option (Yes or No) as chosen in the previous question (p_{stick}). For this model, the likelihood function for the first question is simply equal to p_{Y1} if the Yes insurance option was chosen and $1 - p_{Y1}$ if the No insurance option was chosen; for subsequent questions, the likelihood function is equal to p_{stick} if the option chosen was the same as that chosen by the subject in the previous question of that particular decision problem, and $1 - p_{stick}$ otherwise. The natural log of the likelihood function is summed over all observed choices, and this total log-

Table 13. *Maximum likelihood estimates of the alternative decision models*

Coefficient	Estimate	Standard Error	p-value	Lower 95% Confidence Interval	Upper 95% Confidence Interval	Observations
A. Random Choice model						
p_Y	0.413	0.00570	0.000	0.401	0.424	7452
$H_0 : p_Y = 0.5$			0.000			
B. Random Switching model						
p_{Y1}	0.850	0.0185	0.000	0.814	0.886	7452
p_{stick}	0.849	0.00426	0.000	0.840	0.857	7452
$H_0 : p_{Y1} = 0.5$			0.000			
$H_0 : p_{stick} = 0.5$			0.000			

The log-likelihoods corresponding to the estimations in Panels A and B are -5050.7 and -3167.7. Wald test p-values and 95% confidence intervals reported.

likelihood function can then be maximized with respect to the two parameters (p_{Y1} and p_{stick}) to obtain the maximum likelihood estimates of these parameters. The results of this estimation are reported in Panel B of Table 13.

Panel A of Table 13 shows that the estimated probability of choosing the Yes insurance option in the Random Choice model is approximately 0.41. Additionally, using a Wald test, the hypothesis of $p_Y = 0.5$ is rejected at the 1% significance level, indicating that subjects do not randomize between the two alternatives with equal probability – this is in line with the results of the binomial tests described earlier in this section. For the Random Switching model, the estimates of both p_{Y1} and p_{stick} are around 0.85 (and are significantly different from 0.5). This implies that there is a relatively high estimated probability that participants choose to buy insurance in question 1, which is expected, given that question 1 is associated with the lowest price for the insurance contract (R$0.40). Further, participants also have a relatively high (85%) estimated probability of sticking to their previous choice in questions 2 to 20 – this could explain the observed choice patterns (including the substantial fraction of intransitive choice patterns), in which there are chains of Yes choices and No choices, but there may be multiple switch points or reverse switching (from No to Yes insurance).

To formally test the suitability of the alternative decision models in describing observed choices, I first utilize a Vuong test, developed by Vuong (1989), which is a likelihood ratio test involving the comparison of two non-nested models in order to determine, statistically, which model better describes the data. Vuong tests have been extensively used in the experimental economics literature to compare the fit of non-nested decision models (for example, Loomes et al. 2002, Blavatskyy and Pogrebna 2010, Wilcox 2010). A Vuong test compares pairs of models, say Decision Model 1 and Decision Model 2. The construction of the Vuong test statistic for use in comparing the fit of decision models is described in detail by Loomes et al. (2002); the null hypothesis is that the two models – Decision Models 1 and 2 – are equally close to the true data-generating process (Wilcox 2010). Under this null hypothesis, the test statistic has a limiting standard normal distribution. The test statistic can take on positive or negative values – by construction, positive values of the statistic indicate that Decision Model 1 is closer to the true data-generating process than Decision Model 2, while negative values imply that Decision Model 2 is closer to the true data-generating process (Loomes et al. 2002). A test statistic of 0, on the other hand, implies that both decision models are equally close to the true data-generating process (this is the null hypothesis).

I use Vuong tests comparing each of the two alternative models to each of the six standard models (combined with the Luce error specification involving the free μ parameter) estimated in Section 4.1 – these tests provide an indication of which decision model best fits the experimental data, and thus best describes the choices and latent decision-making process of participants.[86] In each case, the alternative model – that is, Random Choice or Random Switching – is considered to be

[86] I use the standard models combined with, rather than without, the free Luce error (noise) parameter since this combination is more flexible (due to accounting for the extent of errors in decision-making) than the error specification which assumes $\mu = 1$. In addition, I conducted all the estimations with the Fechner error specification, which yield very similar results to the corresponding Luce error estimations.

Table 14. *Results of the Vuong tests for comparing decision models*

Standard Model	Alternative Models	
	Random Choice Model	Random Switching Model
	(1)	(2)
EUT-CRRA + Luce error	-18.41***	20.12***
	(0.000)	(0.000)
EUT-EP + Luce error	-18.40***	20.06***
	(0.000)	(0.000)
PU + Luce error	-22.85***	17.03***
	(0.000)	(0.000)
RDU-CRRA + Luce error	-19.75***	18.96***
	(0.000)	(0.000)
RDU-EP + Luce error	-19.73***	18.97***
	(0.000)	(0.000)
CPT + Luce error	-22.82***	16.99***
	(0.000)	(0.000)
Observations	7452	7452

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Vuong test statistic values reported; p-values corresponding to the one-tailed hypothesis tests are in parentheses.

Decision Model 1, and thus positive values of the Vuong test statistic provide evidence in favour of the alternative model. When the value of the test statistic is positive, I conduct a one-tailed test of the null hypothesis (that is, the test statistic equals 0) against the alternative hypothesis that Decision Model 1 better fits the data (that is, the test statistic is greater than 0); when the test statistic value is negative, I perform a one-tailed test of the null hypothesis against the alternative hypothesis that Decision Model 2 provides a better fit (that is, the test statistic is less than 0). It is important to note that if I fail to reject the null hypothesis using either of these one-tailed tests, I am unable to conclude which of the two models better fits the data – thus, the Vuong test allows for the possibility that neither of the models is "correct" (Wilcox 2010).

The results of these Vuong tests are presented in Table 14. Column (1) of Table 14 presents the Vuong test statistic values (and p-values corresponding to the relevant hypothesis tests) for the comparison of the Random Choice model to the six standard decision models – EUT-CRRA, EUT-EP, PU, RDU-CRRA, RDU-EP and CPT – combined with the Luce error parameter; Column (2) reports the test statistic values and corresponding p-values for the comparison of the Random Switching model to the six standard decision models.[87,88] All the Vuong test statistic values are negative and large in magnitude in Column (1), while they are positive and large in magnitude in Column (2). For all the hypothesis tests in Column (1), the null hypothesis that the Random Choice model and standard decision model considered are equally close to the true data-generating process is convincingly rejected (with a p-value very close to 0), in favour of the alternative hypothesis that the standard decision model better fits the data and is closer to the true data-generating process. On the other hand, for all the hypothesis tests in Column (2), the null hypothesis is also convincingly rejected (with a p-value very close to 0), but in favour of the alternative hypothesis that the Random Switching model has greater explanatory power and is a better candidate for the true data-generating process.[89]

Therefore, the Vuong tests overwhelmingly reject the Random Choice model in favour of the standard decision models, but also indicate that the Random Switching model provides a significantly better fit to the data than the standard models. Additionally, I conducted a Vuong test comparing the Random Switching model (Decision Model 1) to the Random Choice model (Decision Model 2), and obtained a test statistic value of 33.54, implying a convincing rejection of the Random Choice model in favour of the Random Switching model. Thus, the Vuong test results indicate that the Random Switching model performs substantially better than the other decision models in explaining the data, and is the closest – of all the models considered – to the true data-generating process. The Random Switching model is therefore most likely the best candidate to represent the true decision-making process of participants, and the intransitive

[87] As before, the RDU and CPT models utilize the Tvesky-Kahneman probability weighting function.

[88] It is important to note that the Vuong tests, as well as the mixture model estimations, are tests of specific parametric forms of the decision theories, rather than general tests of the theories themselves. Therefore, in this section, I am comparing the suitability of different parametric forms for describing the experimental data.

[89] Note that the Vuong test statistic values obtained here are greater in magnitude than those obtained in most other experimental studies comparing the fit of non-nested decision models.

choice patterns observed in the data are in line with the predictions of this model. These results, combined with those obtained using the constant error rate analysis, indicate that the intransitive choice patterns observed in the data are generated by the true (non-transitive) decision-making process of subjects, given by the Random Switching model, rather than by random errors which cause deviations in observed choices from true preferences that are transitive.[90,91]

The Vuong test, however, implicitly assumes that participants' choices are generated by a single model of behaviour under risk. However, Harrison et al. (2010) and Conte et al. (2011) note that because decision-making under risk is sufficiently heterogenous that it cannot be described by a single theory or model, it is vital to account for the possibility that different theories of decision-making co-exist in the same sample and jointly generate participants' choices – in order to do this, I estimate mixture models of the form specified by Harrison and Rutström (2009) and Harrison et al. (2010). This mixture model estimation is described in detail in Section 5.5 of Chapter 1, and involves the maximization of a grand log-likelihood function in which two decision-making processes co-exist and have different (non-zero) weights; that is, each choice can be generated by either of the two decision models considered in the specification, and the weights indicate the proportion of choices that are better characterized by the corresponding model.[92]

[90] Among the standard decision models considered, CPT (with the free Luce error parameter) performs better in Vuong tests than the five other standard models (with the free Luce error parameter), indicating that the CPT plus error model better fits the data than the other standard decision models. This is along the lines of the result obtained by Galarza (2009), who study data from a MPL experiment conducted in Peru and find that a majority of participants are better characterized by decision-making in accordance with prospect theory. Additionally, Humphrey and Verschoor (2004a) find evidence that the experimental decisions of Ugandan farmers are likely to be best characterized by a stochastic specification of the CPT model; further, they note that both rank dependence and loss aversion are important features of real-world behaviour under risk. Tanaka et al. (2010) use a similar CPT specification to describe the decision-making process under risk of experimental subjects in rural Vietnam (citing evidence from numerous studies which find that such models better describe behaviour under risk than EUT models).

[91] In addition to the Vuong tests reported in Table 14, I also conducted corresponding Clarke tests, and obtained results which are qualitatively similar to – and yield the same conclusions as – those obtained in the Vuong tests. The Clarke test, developed by Clarke (2003) is a directional and symmetric distribution-free test that has also been extensively used for non-nested model comparison and selection.

[92] The mixture model approach assumes that the choices can only be generated by the two latent behavioural processes considered in that particular specification.

Thus, as done for the Vuong tests, I combine each of the two alternative decision models with each of the six standard decision models (in combination with the free Luce error parameter), and use the structural maximum likelihood procedure to obtain the maximum likelihood estimates of the weight parameter. The estimates of this weight parameter provide information on the fraction of choices in the sample that are better characterized by the standard models and the fraction that are better characterized by the two alternative models – thus, this approach is appropriate for comparing the suitability of different non-nested models in describing the data, with larger fractions indicative of a better fit of the corresponding decision model (Harrison and Rutström 2009).

Harrison and Rutström (2009) and Harrison et al. (2010) – who also use such a mixture model approach – obtain maximum likelihood estimates of all the utility function parameters and the weight parameter, by maximizing the grand log-likelihood function jointly with respect to all these parameters. However, when I attempted to maximize (using the data from the Brazilian experiment) the grand log-likelihood functions with respect to all the parameters, the optimization routine did not converge for any combination of standard and alternative decision models. Harrison and Rutström (2009) also point out that one difficulty associated with mixture models is the joint estimation of the weights and the model parameters – if each model has some chance of explaining the data, then mixture models are characterized numerically by relatively flat likelihood functions.[93] This non-convergence associated with the estimation of the mixture model is also encountered in Chapter 1.

Thus, following the strategy used in Chapter 1, I fix the values of the preference parameters and the noise parameter to the maximum likelihood estimates obtained in Section 4.1 for the standard decision models and those obtained earlier in this section for the two alternative models, and maximize the grand log-likelihood function with respect to only the weight parameter. These maximum

[93] However, it is important to note that if the experimental design separates decision rules reasonably strongly, the likelihood function in the mixture model is not expected to be especially flat.

Table 15. *Mixture model estimation results*

Standard Model	Alternative Models	
	Random Choice Model	Random Switching Model
	(1)	(2)
EUT-CRRA + Luce error	0.0830***	0.863***
	(0.0194)	(0.0149)
EUT-EP + Luce error	0.0867***	0.863***
	(0.0193)	(0.0150)
PU + Luce error	0.0172	0.819***
	(0.0177)	(0.0157)
RDU-CRRA + Luce error	0.0644***	0.848***
	(0.0189)	(0.0154)
RDU-EP + Luce error	0.0657***	0.849***
	(0.0189)	(0.0154)
CPT + Luce error	0.0178	0.819***
	(0.0178)	(0.0157)
Observations	7452	7452

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$
Estimates of the weight parameter in the mixture model reported.
Standard errors of these estimates are in parentheses.

likelihood estimates of the weight parameter for the different combinations of standard and alternative models are reported in Table 15, in a table of similar form to that reporting the Vuong test results. In each case, the estimate reported is that of the weight parameter (W) associated with the alternative decision model considered – thus, it provides an estimate of the fraction of choices in the dataset that are better characterized by the alternative model, while the estimated $1 - W$ indicates the proportion of choices better described by the standard decision model considered.[94]

As the results in Column (1) indicate, when the Random Choice model is combined with any of the standard decision models in the mixture specification, the estimate of the weight associated with the Random Choice model is relatively

[94] In accordance with the strategy used by Harrison and Rutström (2009), the maximum likelihood estimation actually provides estimates of the log odds in favor of one model over the other. Denoting the log odds as κ, one can recover the probability of the alternative model being correct as $W = 1/(1+\exp(\kappa))$ (Harrison and Rutström 2009). This non-linear function of κ can be easily calculated from the estimates, and the "delta method" can be used to provide estimates of the standard errors and p-values (Oehlert 1992).

small in magnitude. In each of the six cases, the estimate of the weight parameter is below 0.1, indicating that only a small fraction of choices in the dataset are better characterized by the Random Choice model, while a vast majority (greater than 90%) of the choices are better described by the standard decision model considered in that specification. Additionally, while the estimates of W are statistically significantly different from 0 in four of the mixture model specifications reported in Column (1), the estimate of this parameter when the Random Choice model is combined with the PU or the CPT model is only around 0.017 and is statistically indistinguishable from 0 – in these cases, I cannot reject the hypothesis that none of the choices are better described by the Random Choice model as compared to the standard decision model.

On the other hand, the results in Column (2) show that when the Random Switching model is combined with the standard models, the estimate of the weight parameter (W) is much larger in magnitude. In each of the six cases, the weight parameter is estimated to be greater than 0.8 and statistically significantly different from 0 at the 1% level. These results indicate that more than 80% of the choices in the dataset are better described by the Random Switching model, and thus only a small fraction are better characterized by the standard decision model used in that particular mixture specification.[95] Thus, the results in this table indicate that the Random Switching model performs better than the standard models in describing the observed data, and provides the best representation of the true latent decision-making process of subjects, in line with the results of the Vuong tests.[96] Further, in accordance with the results obtained using the constant error rate analysis and the Vuong tests, the estimates in Table 15 indicate that the inconsistent choice patterns observed in the data are likely to be a result of the true decision-making process of subjects (which is best described by the non-transitive

[95] For all of the weight parameter (W) estimates reported in Table 15, the hypothesis of $W = 0.5$ is convincingly rejected, indicating a rejection of the hypothesis that each of the two decision models considered better explains half of the observed choices.

[96] I also estimated a mixture model combining the Random Switching and Random Choice models, and obtained an estimate of approximately 0.96 for the weight associated with the Random Switching model, implying that around 96% of the observations are better described by this model as compared to the Random Choice model.

Random Switching model), rather than random errors in experimental decision-making (combined with a transitive standard decision model). Meanwhile, the results from both the mixture model estimations and the Vuong tests show that the Random Choice model does not fit the data well.[97]

Thus, the results in this section indicate that the Random Switching model provides a better description of the true decision-making process of participants than any of the standard models or the Random Choice model; in line with the results of Fishburn (1991) and Jacobson and Petrie (2007), I find that the true decision-making process of subjects may be intransitive and it would be inappropriate to assume *a priori* that the decision-making of individuals in the experimental problems is in accordance with one of the standard decision models, or indeed in accordance with any model of behaviour under risk involving well-defined risk preferences (all of which assume transitivity in choice).

Further, while there are heuristics other than that implied by the Random Switching model which could be used to analyze subjects' decisions, and the Random Switching model can fit the data well since it can account for both the intransitive and transitive choice patterns observed, the analysis in this section provides an indication that the true decision-making process of subjects may be represented by such non-transitive decision models which are not based on rational decision-making with well-defined risk preferences.[98] Additionally, Costa-Gomes and Crawford (2006) note (albeit in the context of experimental game theory) that there is an enormous set of possibilities for these decision rules, and it is difficult to identify the exact decision rule(s) used by subjects within this

[97] I conducted the Vuong tests and the mixture model specifications reported in this section separately for the sub-samples of college and non-college educated participants, given that the fractions of inconsistent choices made by these two groups are markedly different (see Section 4.2). For the Vuong tests, the results for both groups are similar to those obtained with the combined sample of all subjects – the Random Choice model is convincingly rejected in favour of the standard decision models, while the standard decision models are convincingly rejected in favour of the Random Switching model. In the mixture model specifications, the estimated fractions of choices best described by the two alternative models (as opposed to the standard decision models) are very similar for both groups and close to the values reported in Table 15 for the combined sample.

[98] In addition, Fishburn (1991) note that describing experimental data using non-transitive models is ultimately an empirical issue, and it is important to provide a good descriptive theory, which may not be normative.

set; in such a situation, it is a viable strategy to analyze a small set of *a priori* plausible rules – in this case, the Random Choice and Random Switching rules – and econometrically test which ones best fit subjects' decisions.[99] Further, Crawford (2007) highlight the importance of studying cognition, cognitive processes and decision rules for obtaining a better understanding of human behaviour and decisions.

If the decision-making process of subjects in the MPL experimental problems follows intransitive decision models, such as the Random Switching model, it indicates that the choices of participants in these decision problems do not reveal true risk preferences, but rather provide information on subjects' preferences about how to play the "games" in a way that is not completely determined by preferences over risk. This is because true rational preferences over risk are transitive, and intransitive choice generated by the true decision-making process of subjects – rather than by random errors – cannot reflect true risk preferences.

The use of such metrics or heuristics, rather than true preferences over risk, to make decisions is likely due to poor understanding of the decision problems, and caused the relatively large fraction of intransitive choice patterns observed in the data. Further, the analysis in the previous section indicated that the intransitive choice patterns are not likely to be caused by indifference between alternatives, while the constant error rate analysis indicated that they are also unlikely to be caused by random errors. Thus, the results in this chapter suggest that the intransitive choice patterns observed in the Brazilian experiment are most likely

[99] It is important to note that I do not claim that the Random Switching model (or heuristic) is the best model to describe the data – there may be other models or heuristics that fit the data better. However, I wanted to test the suitability of a restrictive model – which includes just one or two parameters (such as the Random Switching model) – for describing the data, and test whether such a restrictive model model not based on well-defined preferences over risk could describe the data better than the standard decision models considered (that is, the EUT, RDU and CPT models analyzed in this chapter). I did find this to be the case, but the Random Switching model is just one possible decision rule – which I considered to be *a priori* plausible for describing decisions in the MPL problems – that performs better than the standard models, and there are likely to be others. However, since this restrictive Random Switching model beats the standard decision models, I can reject these standard models in favour of the Random Switching model for describing the experimental data at hand, and there is an indication that the true decision-making process of subjects may be represented by such non-transitive decision models which are not based on rational decision-making with well-defined risk preferences.

due to poor comprehension of the decision problems, which in turn is likely a consequence of the complexity of the MPL elicitation procedure.[100,101]

Humphrey and Verschoor (2004a) also note that experiment participants may utilize simplifying heuristics in decision-making, and report choices that are unrelated to true preferences, when they face difficulty in comprehending decision problems. Additionally, Dave et al. (2010) find that the prevalence of inconsistent choice patterns is likely due to poor understanding of the complex MPL elicitation procedure (which places a greater "cognitive load" on subjects); Charness and Viceisza (2012) reach a similar conclusion in their analysis of experimental data from rural Senegal. Thus, the results in this chapter indicate that the elicited responses of subjects may not reflect true risk preferences, primarily because proper understanding of the decision problems is crucial for eliciting meaningful responses that accurately reflect true risk preferences

[100] For the 122 observed inconsistent choice patterns involving multiple switching, the average difference between the first-switch question number and the last-switch question number is approximately 10 – the risk preferences implied by such widely separated switch points are markedly different, further indicating a lack of understanding of the decision problems.

[101] Dave et al. (2010) find that when participants in their experiment did not understand the MPL decision problem, they tended to resort to alternating between the two options, describing their choices with statements such as "I just wanted to try it".

(Humphrey and Verschoor 2004a, Charness et al. 2013).[102,103,104] Additionally, this implies that the preference parameter estimates of the different decision models (obtained in Section 4.1) may not provide an accurate characterization of subjects' risk attitudes, and caution must be exerted in interpreting these estimates.

Thus, the results indicate that the MPL procedure, which is "a mainstay of risk elicitation in experimental economics" (and, to date, the most popular risk elicitation procedure, as noted by Crosetto and Filippin 2013), is not suitable for eliciting consistent choices that accurately reflect true risk preferences in developing country settings (and is possibly also unsuitable for non-standard

[102] Barr and Packard (2005) and Charness and Viceisza (2012) note that experimental decision problems which are framed in agricultural terms (for example, in terms of "yields" and "seeds"), as compared to those framed in the abstract, generally produce higher rates of comprehension in rural regions of developing countries, since most subjects in these areas can better relate to notions of risk in agricultural terms. Thus, the abstract framing used in the MPL decision problems of the Brazilian experiment could also have contributed to the poor understanding and observed inconsistencies in choice. However, to comprehensively test the hypothesis that framing decision problems as real-life ones increases participant understanding, an experiment is required which includes decision problems which are similar in every way except that some are framed in the abstract and some are framed in real-world agricultural terms.

[103] It is unlikely that a lack of financial motivation is the cause of the randomness in observed choice – as noted in Section 3.1, the monetary prizes in the Brazilian experiment were high enough (relative to average daily income) to provide sufficient financial incentive for participants to answer truthfully in the decision problems (Falco 2012).

[104] It may also be the case that the complex nature of the insurance contracts – particularly the index insurance contracts – offered in the decision problems of the Brazilian experiment is also a cause of the poor understanding. However, in Chapter 4 I find that understanding and cognitive ability do not affect the take-up of index insurance contracts in an Ordered Lottery Selection experiment conducted in rural Ethiopia. Additionally, in line with this observation, Hill and Nobles (2011) find that weather indexed insurance products offered to farmers in Southern Ethiopia in both an experimental setting and in real life were well understood, and in particular the participants involved were able to comprehend the basis risk associated with the policies – as a result, they note that such products may be appropriate for populations with relatively low quantitative and financial literacy. Similarly, Hill et al. (2011) also find evidence that most participants in the Ethiopian Rural Household Survey understood a hypothetical index insurance product offered, as well as the associated basis risk. Given that the Brazilian sample is relatively highly educated – and thus subjects in this sample are expected to have greater exposure to cognitive exercises – as compared to inhabitants of rural Ethiopia who have limited access to formal education, it is expected that the participants in the Brazilian experiment appropriately understood the nature of the insurance contracts offered. Thus, the complex MPL elicitation procedure used in the Brazilian experiment is likely to be the major cause of the poor comprehension.

experimental samples in developed countries).[105] Rather, simpler methods, such as the Ordered Lottery Selection procedure, may be preferred in these settings, since the existing literature on the topic indicates that these methods are more likely to accurately elicit true risk preferences, for a larger set of individuals.[106]

5 Conclusion

In this chapter, I analyze a wide variety of decision models using experimental data from Brazil. Unlike the Ethiopian experiment considered in Chapters 1 and 2, the Brazilian experiment analyzed in this chapter used the Multiple Price List elicitation procedure, in which participants made a series of 20 binary choices in each decision problem (as opposed to a single choice in each Ordered Lottery Selection decision problem of the Ethiopian experiment). I find that a wider range of decision models can be estimated using the Brazilian MPL data as compared to when using the Ordered Lottery Selection data from the Ethiopian experiment, primarily due to the greater precision and power provided by the MPL elicitation procedure (in line with the observation of Dave et al. 2010).

In particular, I estimate six established and popular "standard" decision models – EUT-CRRA, EUT-EP, PU, RDU-CRRA, RDU-EUP and CPT – using the Brazilian data, which enables an extensive and comprehensive analysis of the risk preferences of experimental subjects. The estimates of most preference parameters (for example, the probability weighting parameter) differ significantly when using data from the Brazilian and Ethiopian experiments – this implies quantitatively and qualitatively different behaviour under risk for the two samples

[105] Among college-educated subjects in the Brazilian sample, approximately 25% of choices are intransitive. While this is substantially less than the fraction of intransitive choices for non-college educated subjects in the sample (approximately 55%), it is still significantly greater than that observed in most MPL experiments conducted with standard samples of university students in developed countries (around 10%). Therefore, further research to determine whether the MPL procedure is suitable for samples of university students in developing countries is essential.

[106] However, it is important to acknowledge that since the true underlying risk preferences of individuals are not known, it is difficult to gauge whether elicited risk attitudes accurately reflect true risk preferences (Dave et al. 2010).

of subjects, highlighting the difficulty in extending experimental results from one setting to another. However, the estimation of risk preferences of individuals in a particular region is crucial, as the design and evaluation of policies involving risky outcomes – such as those relating to the adoption of new technologies or seed varieties – should appropriately account for individuals' risk attitudes (Harrison 2011).

I also estimate the standard decision models in conjunction with two popular error specifications, the Luce and Fechner error models, which account for the magnitude of noise in decision-making. The results of these estimations imply a considerable degree of randomness in participant choice. Additionally, I find that a substantial fraction (33%) of choice patterns in the decision problems are intransitive – due to multiple or reverse switching – and thus not consistent with any of the standard decision models, and indeed with any well-defined preferences over risk. Barr (2007) notes that the presence of such inconsistent choice patterns in MPL decision problems could indicate a lack of understanding, random errors in decision-making or genuine indifference between alternatives. An important aim of this chapter is to shed light on which of these factors is the most likely cause of the inconsistent patterns observed in the Brazilian experiment.

In Probit and Ordered Probit regressions, I find that participants who did not attend college are significantly more likely to have intransitive choice patterns, and also to make a larger number of inconsistent reverse switches (from the No insurance to the Yes insurance option), indicating that inconsistencies in choice could be related to low exposure to cognitive exercises and low cognitive ability. Similarly, when allowing for heterogeneity in the noise parameter, I also find that lower education is related to a larger noise parameter estimate in the structural maximum likelihood estimations of the Luce and Fechner error models.

In the estimations that account for heterogeneity in the noise parameter, the preference parameters of the decision models are assumed to be constant across subjects – this enables the comparison of the extent of noise in decision-making through a direct comparison of the estimated noise parameter. However, the results of studies (for example, Dave et al. 2010, Crosetto and Filippin 2013) that utilize a comparison of noise parameter estimates to compare the extent of errors in decision-making across contexts in which the preference parameters differ could be misleading, since the estimate of the noise parameter depends also on the preference parameter values, and variation in the noise parameter does not solely capture differences in noise. Therefore, in this chapter, I propose a normalization procedure which enables the direct comparison of the noise parameter across contexts in which preference parameters differ. Comparing the normalized noise parameter estimates across the Brazilian and Ethiopian experiments analyzed in this thesis, I find that the extent of errors in decision-making is greater in the Brazilian experiment, likely due to poor understanding of the complex MPL decision problems contained in this experiment.

Further, using the constant error rate analysis (developed by Harless and Camerer 1994), I find evidence that the inconsistent choice patterns observed in the data are representative of the true decision-making process (which is non-transitive) rather than random (or "trembling hand") errors. In order to further explore this possibility, I estimate two alternative decision models – the Random Switching and the Random Choice models – which, unlike the standard models, do not assume transitivity, and evaluate how well they describe observed choices.

Using Vuong tests and mixture model estimations, I find that while the Random Choice model does not fit the data well, the Random Switching model, in which participants choose a particular option in the first binary question of a decision problem with a particular (fixed) probability and then in subsequent questions stick to their previous choice with another (fixed) probability, better fits the data than any of the standard models considered. Indeed, the results indicate that the Random Switching model is the most suitable candidate – of all the models considered – for representing the true latent decision-making process of participants; this implies that it may not be appropriate for researchers to implicitly assume transitivity in experimental choice. Further, since the Random Switching model is not based on rational decision-making with well-defined risk preferences and generates intransitive choice patterns, the results indicate that the choices of participants in the MPL decision problems do not reveal true risk preferences (which are presumably transitive), but rather provide information on subjects' preferences about how to play the "games" in a way that is not completely determined by preferences over risk. This reduces the reliability of risk preference parameter estimates, and caution must be exercised when interpreting results from MPL experiments which have similarly high fractions of intransitive choices, as observed choices and elicited risk attitudes could diverge substantially from true preferences over risk.[107] Additionally, I find that indifference between alternatives in the decision problems is unlikely to be a cause of the multiple switching patterns observed in the data.

Thus, it is likely that participants use the metric or heuristic implied by the Random Switching model for making decisions due to confusion and poor understanding of the decision problems. This once again highlights the issue of poor understanding associated with MPL experiments conducted in developing countries, and poor comprehension is the most likely cause of the large fraction of inconsistent choice patterns observed in the Brazilian data, rather than indifference or "trembling hand" errors. Thus, the results in this chapter – along the lines of those obtained by Charness and Viceisza (2012) – indicate that the complex MPL method, which has been used extensively in both developed and developing countries, may not be suitable for experiments conducted with poorly-educated subjects in developing country settings. For these settings, as noted by Charness et al. (2013), single-choice tasks (such as Ordered Lottery Selection problems) should be preferred to multiple-choice tasks, since the advantages of

[107] It is important to note that the findings in this chapter do not imply that the external validity of experimental results, in general, is extremely limited. The findings do not provide a general comment about all experiments, but only extend to experiments in which the decision problems are not well understood by participants and thus choices do not reflect true preferences over risk. For such experiments, the external validity of findings is limited, and assuming that observed choices reflect true preferences is misleading.

complex methods – primarily greater precision and power to estimate different decision models – are outweighed by the disadvantage of lower comprehension (which leads to spurious responses that do not reflect true risk preferences). Researchers using the MPL procedure with poorly-educated subjects should take serious consideration of this issue, and include procedures to check and ensure participant understanding.[108]

An attractive alternative to both the MPL and Ordered Lottery Selection methods for experiments conducted in developing countries is the simple, single-choice investment task used by Gneezy and Potters (1997) to elicit risk preferences. In such a task, each experiment participant is provided with an initial endowment and is required to choose how much of this endowment he would prefer to invest in a risky asset with a positive expected profit; the amount that is not invested is kept by the participant. It is possible to directly compute the coefficient of relative risk aversion from the investment choice of a subject (Charness and Viceisza 2012). Furthermore, Charness and Viceisza (2012) note that this task is relatively easy to comprehend, and find that it is well understood by a sample of farmers in rural Senegal (who have low levels of formal education). The Gneezy and Potters (1997) experimental task has been widely implemented in developing countries (for example, Gneezy et al. 2009, Gong and Yang 2012).

The results in this chapter highlight the importance of choosing the right experimental procedure – with the appropriate level of complexity – for obtaining accurate estimates of risk attitudes (Dave et al. 2010). Dave et al. (2010) and Charness et al. (2013) conclude that it is imperative for researchers to consider the tradeoffs – primarily between precision and comprehension – for different elicitation methods to determine the most suitable alternative; further, they note that the choice of task should be dependent on the setting and the characteristics of the sample population, and warn that a task suitable for one setting may not be appropriate for accurately eliciting risk preferences in another setting.

[108] For example, detailed and lengthy learning rounds, as used by Lybbert et al. (2010), conducted prior to the main decision problems could aid and improve participant understanding of the experimental procedures.

However, Charness and Viceisza (2012) note that this issue is often ignored by economists and, in particular, there is a distinct lack of experimental studies comparing the effectiveness of different elicitation procedures in developing countries. Such a study should involve an experiment which includes decision problems using different elicitation procedures (such as the Multiple Price List and Ordered Lottery Selection methods), conducted with the same sample of developing country participants (along the lines of Charness and Viceisza 2012) – this would enable a direct comparison of the effectiveness of different procedures for accurately eliciting risk preferences.[109]

A promising avenue for further research is to explore the possibility of creating variations of the MPL design that are easily understood. For example, using a MPL variant in which participants are presented choices in a visual, one-at-a-time format, Eckel et al. (2007) find a significant increase in comprehension and consistency. Thus, the Eckel et al. (2007) study shows that the MPL procedure can be implemented in a way that is understood by poorly-educated subjects in a developing country. Given the greater precision associated with the MPL procedure, improving the implementation and increasing the comprehension of the MPL method is preferred to abandoning it in favour of the Ordered Lottery Selection procedure, for experiments conducted in developing countries.

Alternatively, the possibility of increasing the precision of the simpler Ordered Lottery Selection design should also be explored. For example, an analogue of the Iterative MPL procedure (outlined by Andersen et al. 2006b) could be considered for Ordered Lottery Selection problems. Such a procedure would include a series of iterative Ordered Lottery Selection problems presented to subjects – the first would include options that cover a wide range of risk preferences; after a subject's choice in the first problem, the second problem presented would allow the subject to make a choice from refined options within the last option chosen (Andersen et al. 2006b). Since the second problem contains options which only

[109] Further, Crosetto and Filippin (2013) argue that it is preferable to conduct a comparative study of this type using a between-subject, rather than a within-subject, analysis.

allow for risk preferences within the interval implied by the participant's first choice, this procedure enables a finer classification of risk attitudes as compared to a single Ordered Lottery Selection problem, while preserving the simplicity of the Ordered Lottery Selection design. This process could be continued until the desired level of precision in elicited risk preferences is obtained.

As noted by Jacobson and Petrie (2009) and Hirschauer et al. (2012), most experimental studies ignore, or entirely discard, inconsistent choices in order to avoid biases in the interpretation of MPL data, under the assumption that these choices represent uninformative noise. However, the analysis conducted in Section 4.3 highlights the importance of considering inconsistencies in choice for obtaining a better understanding of experimental decision-making. Further, Charness and Viceisza (2012) note that the removal of a large fraction of inconsistent choices could lead to skewed results (which may also not be appropriate for making policy recommendations); this is particularly true when the probability of making an inconsistent choice is related to the education level or cognitive ability of a subject (Charness et al. 2013).

Variations of the MPL design have been created which impose consistency in choice – such "Switching MPL" procedures enforce a single switch point by simply asking subjects to choose the question number (or row) at which they would switch from the safe to the risky option; the other non-switch choices are then inferred automatically, imposing monotonicity on revealed preferences and enforcing transitivity (Charness et al. 2013). Under these assumptions, the risk preferences of a participant can be inferred from his chosen switch point (Andersen et al. 2006b). Such variations of the MPL design have been used, for example, by Andersen et al. (2006b), Liu (2008) and Tanaka et al. (2010). This design allows the researcher to bypass issues relating to inconsistencies in subject choice; additionally, Andersen et al. (2006b), analyzing experimental data from Denmark and the United States, find no significant difference in risk attitudes elicited using the Switching MPL and original MPL procedures, concluding that "nothing is lost from using an enforced single switch point, and that the (multiple) switching behaviour that is often observed in MPL simply reflects indifference".

However, this finding may not hold for experiments conducted in developing countries – as the results of this chapter indicate, the assumption of consistency or transitivity in experimental choice may be inaccurate in such settings.[110] Subjects – particularly those with low levels of formal education, who characterize developing country experimental samples – may face difficulty in understanding the complex MPL design in a Switching MPL experiment, and thus make choices in a psuedo-random fashion which are consistent by design, but do not reflect or reveal true preferences over risk. Therefore, it is likely that the issue of poor comprehension also arises in Switching MPL experiments, particularly those conducted in developing countries by, for example, Liu (2008) (rural China) and Tanaka et al. (2010) (rural Vietnam). It is possible that participants in these experiments choose their single switch point at random or in a manner that does not reflect their true risk attitudes – if these responses are assumed to reflect true preferences over risk, it would significantly bias the estimates of preference parameters and the results of these experimental studies could be severely misleading. Further, Liu (2008) and Tanaka et al. (2010) assume deterministic choices and do not account for the (strong) possibility that subjects make errors in decision-making – with the deterministic choice assumption, relatively few choices are required to identify preference parameters, but the parameter values obtained may not accurately reflect true risk attitudes. An important avenue for further research is to determine how misleading the results of these studies (which use single-switch variations of the MPL procedure) may be if randomness in choice was commensurate to what is found in standard MPL experiments conducted in similar contexts. Additionally, it is essential to explore how such an experimental design could be improved by adding choices that

[110] The possible invalidity of this assumption is also highlighted by the extensive evidence indicating that a significant fraction of subjects make inconsistent choices in MPL experiments conducted both in developed and developing countries. Charness et al. (2013) also draw attention to the possibility that experiments using the Switching MPL procedure impose added assumptions on risk preferences which may not hold and could lead to biased results. Further, Harrison and Rutström (2008) note that obtaining a sample of consistent choices should not be the sole criterion for choosing a particular elicitation procedure, especially if a stochastic choice process is allowed for.

help detect, and test for, inconsistencies in choice, errors in decision-making and evidence of poor comprehension.

Furthermore, Jacobson and Petrie (2009), using data from Rwanda, find that inconsistencies in experimental choice are linked to economic decisions in real life. In particular, they find that individuals who make inconsistent choices in their experiment forego potentially beneficial opportunities and make sub-optimal financial decisions outside the experimental setting. The fact that the subjects involved in the Brazilian experiment routinely make decisions regarding insurance purchase – which have a notable impact on welfare – emphasizes the importance of understanding their decision-making process under risk and accounting for inconsistent choices, which could be occurring in these crucial decisions outside the experiment as well. Thus, given the potential importance of inconsistent choices, further research on the correlates of these choices, as well as on the relationship between inconsistencies in experimental choice and economic decisions in real life, is vital.

REFERENCES

Andersen, S., G.W. Harrison, and E.E. Rutström, "Choice behavior, asset integration and natural reference points," *Department of Economics Working Papers*, 2006.

——, ——, **M.I. Lau, and E.E. Rutström**, "Elicitation using multiple price list formats," *Experimental Economics*, 2006, *9* (4), 383–405.

Andreoni, James and Charles Sprenger, "Estimating Time Preferences from Convex Budgets," *The American Economic Review*, 2012, *102* (7), 3333–3356.

Barr, Abigail, "Attitudes to Risk in Ghana: Field Manual," 2007.

—— **and Truman G Packard**, "Seeking solutions to Vulnerability in Old Age: Preferences, constraints, and alternatives for coverage under PeruÕs pension system," *The Centre for the Study of African Economies Working Paper Series*, 2005, p. 239.

Barros, A.J.D. and C.G. Victora, "A nationwide wealth score based on the 2000 Brazilian demographic census," *Revista de Saúde Pública*, 2005, *39* (4), 523–529.

Binswanger, H.P., "Attitudes toward risk: Experimental measurement in rural India," *American Journal of Agricultural Economics*, 1980, *62* (3), 395–407.

Blavatskyy, P.R. and G. Pogrebna, "Models of stochastic choice and decision theories: why both are important for analyzing decisions," *Journal of Applied Econometrics*, 2010, *25* (6), 963–986.

Broad, K., A. Pfaff, R. Taddei, A. Sankarasubramanian, U. Lall, and F. de Assis de Souza Filho, "Climate, stream flow prediction and water management in northeast Brazil: societal trends and forecast value," *Climatic Change*, 2007, *84* (2), 217–239.

Camerer, C.F. and T.H. Ho, "Violations of the betweenness axiom and nonlinearity in probability," *Journal of Risk and Uncertainty*, 1994, *8* (2), 167–196.

Ceci, Stephen J, "How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence.," *Developmental psychology*, 1991, *27* (5), 703.

Cerny, Ales, *Mathematical techniques in finance: tools for incomplete markets*, Princeton University Press, 2009.

Charness, G. and A. Viceisza, "Comprehension and risk elicitation in the field: Evidence from rural Senegal," *IFPRI Discussion Paper*, 2012.

Charness, Gary, Uri Gneezy, and Alex Imas, "Experiential methods: Eliciting risk preferences," *Journal of Economic Behavior & Organization*, 2013.

Clarke, K.A., "Nonparametric model discrimination in international relations," *Journal of Conflict Resolution*, 2003, *47* (1), 72.

Coller, M. and M.B. Williams, "Eliciting individual discount rates," *Experimental Economics*, 1999, *2* (2), 107–127.

——, **G.W. Harrison, and E.E. Rutström**, "Are discount rates constant? Reconciling theory and observation," Technical Report 2003.

Conte, Anna, John D Hey, and Peter G Moffatt, "Mixture models of choice under risk," *Journal of Econometrics*, 2011, *162* (1), 79–88.

Costa-Gomes, MA and VP Crawford, "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study," *American Economic Review*, 2006, *96* (5), 1737–1768.

Crawford, Vincent P, "Look-ups as the Windows of the Strategic Soul: Studying Cognition via Information Search in Game Experiments," Technical Report, UCLA Department of Economics Working Paper 2007.

Crawford, V.P. and J. Meng, "New York City Cab Drivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational Expectations Targets for Hours and Income," *The American Economic Review*, 2011, *101* (5), 1912–1932.

Crosetto, Paolo and Antonio Filippin, "A Theoretical and Experimental Appraisal of Five Risk Elicitation Methods," Technical Report, Friedrich-Schiller-University Jena, Max Planck Institute of Economics 2013.

Dave, C., C.C. Eckel, C.A. Johnson, and C. Rojas, "Eliciting risk preferences: When is simple better?," *Journal of Risk and Uncertainty*, 2010, pp. 1–25.

Doerr, U., T. Mahmoud, and U. Schmidt, "Overconfidence and Risk Management of Ethiopian Farmers," Technical Report, Working Paper 2011.

Eckel, Catherine, Jim Engle-Warnick, and Cathleen Johnson, "Adaptive Elicitation of Risk Preferences," *University of Montreal Working Paper*, 2007.

Epley, Nicholas and Thomas Gilovich, "The anchoring-and-adjustment heuristic Why the adjustments are insufficient," *Psychological Science*, 2006, *17* (4), 311–318.

Fafchamps, Marcel, "Vulnerability, Risk Management, and Agricultural Development," *CSAE Working Paper Series*, 2009.

Falco, P., "Does risk matter for occupational choices? Experimental Evidence from an African labour market," *Manuscript submitted for publication*, 2012.

Falk, Armin and Markus Knell, "Choosing the Joneses: Endogenous goals and reference standards," *The Scandinavian Journal of Economics*, 2004, *106* (3), 417–435.

Farber, H.S., "Reference-dependent preferences and labor supply: The case of New York City taxi drivers," *The American Economic Review*, 2008, *98* (3), 1069–1082.

Fechner, G., "Elements of psychophysics. Vol. I.," 1966.

Ferreira, F.H.G. and P. Lanjouw, "Rural nonfarm activities and poverty in the Brazilian Northeast," *World Development*, 2001, *29* (3), 509–528.

Filmer, D. and L.H. Pritchett, "Estimating Wealth Effects Without Expenditure Data – Or Tears: An Application To Educational Enrollments In States Of India," *Demography*, 2001, *38* (1), 115–132.

Fishburn, P.C., "Nontransitive preferences in decision theory," *Journal of Risk and Uncertainty*, 1991, *4* (2), 113–134.

Galarza, F., "Choices under risk in rural Peru," *Munich RePec Archive Working Paper 17708*, 2009.

Gneezy, U. and J. Potters, "An experiment on risk taking and evaluation periods," *The Quarterly Journal of Economics*, 1997, *112* (2), 631–645.

Gneezy, Uri, Kenneth L Leonard, and John A List, "Gender differences in competition: Evidence from a matrilineal and a patriarchal society," *Econometrica*, 2009, *77* (5), 1637–1664.

Gong, Binglin and Chun-Lei Yang, "Gender differences in risk attitudes: Field experiments on the matrilineal Mosuo and the patriarchal Yi," *Journal of Economic Behavior & Organization*, 2012, *83* (1), 59–65.

Gonzalez, R. and G. Wu, "On the shape of the probability weighting function," *Cognitive Psychology*, 1999, *38* (1), 129–166.

Harless, D.W. and C.F. Camerer, "The predictive utility of generalized expected utility theories," *Econometrica: Journal of the Econometric Society*, 1994, pp. 1251–1289.

Harrison, G.W., "Experimental methods and the welfare evaluation of policy lotteries," *European Review of Agricultural Economics*, 2011, *38* (3), 335–360.

____ **and E.E. Rutström**, "Risk aversion in the laboratory," in J. C. Cox and G. W. Harrison, eds., *Research in Experimental Economics*, Vol. 12, Emerald Group Publishing Limited, 2008, pp. 41–196.

____ **and** ____ , "Expected utility theory and prospect theory: One wedding and a decent funeral," *Experimental Economics*, 2009, *12* (2), 133–158.

____ **and J.A. List**, "Field experiments," *Journal of Economic Literature*, 2004, *42* (4), 1009–1055.

____ **, E. Johnson, M.M. McInnes, and E.E. Rutström**, "Individual choice and risk aversion in the laboratory: A reconsideration," *University of Central Florida, Department of Economics Working Paper*, 2003, (3-18).

____ **, J.A. List, and C. Towe**, "Naturally occurring preferences and exogenous laboratory experiments: A case study of risk aversion," *Econometrica*, 2007, *75* (2), 433–458.

____ **, M.I. Lau, and E.E. Rutström**, "Theory, experimental design and econometrics are complementary (and so are lab and field experiments)," *The Methods of Modern Experimental Economics*, 2011.

____ **,** ____ **, and M.B. Williams**, "Estimating individual discount rates in Denmark: A field experiment," *The American Economic Review*, 2002, *92* (5), 1606–1617.

____ **,** ____ **, E.E. Rutström, and M.B. Sullivan**, "Eliciting risk and time preferences using field experiments: Some methodological issues," *Field experiments in economics*, 2005, *10*, 337–372.

____ **, S.J. Humphrey, and A. Verschoor**, "Choice under Uncertainty: Evidence from Ethiopia, India and Uganda," *The Economic Journal*, 2010, *120* (543), 80–104.

Heinemann, F., "Measuring risk aversion and the wealth effect," *Research in Experimental Economics*, 2008, *12*, 293–313.

Hey, J.D., "Experimental investigations of errors in decision making under risk," *European Economic Review*, 1995, *39* (3-4), 633–640.

____ **and C. Orme**, "Investigating generalizations of expected utility theory using experimental data," *Econometrica: Journal of the Econometric Society*, 1994, pp. 1291–1326.

____ **and E. Carbone**, "Stochastic choice with deterministic preferences: An experimental investigation," *Economics Letters*, 1995, *47* (2), 161–167.

Hill, R. V. and M. Nobles, "Flexible insurance for heterogenous farmers: results from a small scale pilot in Ethiopia," *IFPRI discussion paper 01092*, 2011.

Hill, R.V., J. Hoddinott, and N. Kumar, "Adoption of weather index insurance: Learning from willingness to pay among a panel of households in rural Ethiopia," *IFPRI discussion paper 01088*, 2011.

Hirschauer, N., S. Maart, and O. Musshoff, "How to avoid mean and variance bias in Hold-and-Laury lotteries – a note," *Working Paper*, 2012.

Holt, C.A. and S.K. Laury, "Risk aversion and incentive effects," *American Economic Review*, 2002, *92* (5), 1644–1655.

Howe, L.D., J.R. Hargreaves, and S.R.A Huttly, "Issues in the Construction of Wealth Indices for the Measurement of Socio-Economic Position in Low-Income Countries," *Emerging Themes in Epidemiology*, 2008, *5* (3).

Humphrey, S.J. and A. Verschoor, "Decision-making under risk among small farmers in East Uganda," *Journal of African Economies*, 2004, *13* (1), 44.

____ **and** ____ , "The probability weighting function: experimental evidence from Uganda, India and Ethiopia," *Economics Letters*, 2004, *84* (3), 419–425.

Humphrey, Steven J, "Do anchoring effects underlie event-splitting effects? An experimental test," *Economics Letters*, 1996, *51* (3), 303–308.

Jacobson, S. and R. Petrie, "Inconsistent Choices in Lottery Experiments: Evidence from Rwanda," *Experimental Economics Center Working Paper Series*, 2007.

____ **and** ____ , "Learning from mistakes: What do inconsistent choices over risk tell us?," *Journal of Risk and Uncertainty*, 2009, *38* (2), 143–158.

Jacoby, H.G. and E. Skoufias, "Risk, financial markets, and human capital in a developing country," *The Review of Economic Studies*, 1997, pp. 311–335.

Kahneman, D., J.L. Knetsch, and R.H. Thaler, "Experimental tests of the endowment effect and the Coase theorem," *Journal of Political Economy*, 1990, pp. 1325–1348.

Laury, S.K. and C.A. Holt, "Further Reflections on Prospect Theory," *Experimental Economics Center Working Paper Series*, 2005.

Liu, Elaine M, "Time to change what to sow: Risk preferences and technology adoption decisions of cotton farmers in China," *Review of Economics and Statistics*, 2008.

Loomes, G., "Modelling the stochastic component of behaviour in experiments: Some issues for the interpretation of data," *Experimental Economics*, 2005, *8* (4), 301–323.

——, **P.G. Moffatt, and R. Sugden**, "A microeconometric test of alternative stochastic theories of risky choice," *Journal of risk and Uncertainty*, 2002, *24* (2), 103–130.

Luce, R.D., "Individual choice behavior.," 1959.

Lusk, J.L. and K.H. Coble, "Risk aversion in the presence of background risk: Evidence from an economic experiment," *Research in Experimental Economics*, 2008, *12*, 315–340.

Lybbert, Travis J, Francisco B Galarza, John McPeak, Christopher B Barrett, Stephen R Boucher, Michael R Carter, Sommarat Chantarat, Aziz Fadlaoui, and Andrew Mude, "Dynamic Field Experiments in Development Economics: Risk Valuation in Morocco, Kenya, and Peru," *Agricultural and Resource Economics Review*, 2010, *39*, 1–17.

Meier, S. and C. Sprenger, "Impatience and credit behavior: evidence from a field experiment," *Federal Reserve Bank of Boston Working Paper Series*, 2007.

Moser, C. and A. Felton, "The construction of an asset index measuring asset accumulation in Ecuador," *Chronic Poverty Research Centre Working Paper*, 2007.

Munro, Alistair and Robert Sugden, "On the theory of reference-dependent preferences," *Journal of Economic Behavior & Organization*, 2003, *50* (4), 407–428.

Oehlert, G.W., "A note on the delta method," *American Statistician*, 1992, *46* (1), 27–29.

Pfaff, A. and M.A. Vélez, "Efficiency and Equity in Negotiated Resource Transfers: contributions and limitations of trust with limited contracts," *Ecological Economics*, 2011.

——, ——, **R. Taddei, and K. Broad**, "Unequal Information, Unequal Allocation: Bargaining field experiments in NE Brazil," *Environmental Science & Policy*, 2012.

Prasad, K. and T.C. Salmon, "Self selection and market power in risk sharing contracts," Technical Report 2007.

Prelec, D., "The probability weighting function," *Econometrica*, 1998, *66* (3), 497–527.

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

——, *Generalized expected utility theory: The rank-dependent model*, Springer, 1993.

Rabin, M., "Risk aversion and expected-utility theory: A calibration theorem," *Econometrica*, 2000, *68* (5), 1281–1292.

Saha, A., "Expo-power utility: A'flexible' form for absolute and relative risk aversion," *American Journal of Agricultural Economics*, 1993, pp. 905–913.

Sakia, RM, "The Box-Cox transformation technique: a review," *The statistician*, 1992, pp. 169–178.

Schellenberg, J.A., C.G. Victora, A. Mushi, D. de Savigny, D. Schellenberg, H. Mshinda, and J. Bryce, "Inequities among the very poor: health care for children in rural southern Tanzania," *The Lancet*, 2003, *361* (9357), 561–566.

Schmidt, Ulrich, "Efficient risk-sharing and the dual theory of choice under risk," *The Journal of Risk and Insurance*, 1999, *66* (4), 597–608.

Segal, Uzi and Avia Spivak, "First order versus second order risk aversion," *Journal of Economic Theory*, 1990, *51* (1), 111–125.

Sopher, B. and G. Gigliotti, "Intransitive cycles: Rational Choice or random error? An answer based on estimation of error rates with experimental data," *Theory and decision*, 1993, *35* (3), 311–336.

Starmer, C., "Cycling with rules of thumb: An experimental test for a new form of non-transitive behaviour," *Theory and Decision*, 1999, *46* (2), 139–157.

——, "Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk," *Journal of Economic Literature*, 2000, *38* (2), 332–382.

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

Tanaka, Y. and A. Munro, "Regional Variation in Risk and Time Preferences: Evidence from a Large-Scale Field Experiment in Rural Uganda," *National Graduate Institute for Policy Studies Working Paper, Tokyo*, 2012.

Train, K., *Discrete choice methods with simulation*, second ed., Cambridge Univ Pr, 2009.

Tversky, A. and D. Kahneman, "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and uncertainty*, 1992, *5* (4), 297–323.

Tversky, Amos and Daniel Kahneman, "Judgment under Uncertainty: Heuristics and Biases," *Science*, 1974, *185* (415), 1124–1131.

Vuong, Q.H., "Likelihood ratio tests for model selection and non-nested hypotheses," *Econometrica: Journal of the Econometric Society*, 1989, *57* (2), 307–333.

Vyas, S. and L. Kumaranayake, "Constructing socio-economic status indices: how to use principal components analysis," *Health Policy and Planning*, 2006, *21* (6), 459–468.

Wilcox, N.T., "Stochastic models for binary discrete choice under risk: A critical primer and econometric comparison," 2008.

——, "A Comparison of Three Probabilistic Models of Binary Discrete Choice Under Risk," *Working Paper*, 2010.

Appendix A: Estimation Results for the Constant Error Rate Analysis

Table 16. *Maximum likelihood estimates of constant error rate model*

Parameter	Restricted Model (1)	Unrestricted Model (2)
P_2	0.0320**	0.0318***
	(0.0133)	(0.0123)
P_3	0.0573***	0.0493***
	(0.0158)	(0.0150)
P_4	0.0580***	0.0488***
	(0.0165)	(0.0149)
P_5	0.0288**	0.0310**
	(0.0127)	(0.0133)
P_6	0.127***	0.121***
	(0.0207)	(0.0184)
P_7	0.0306**	0.0318**
	(0.0130)	(0.0126)
P_8	0.0459***	0.0440***
	(0.0140)	(0.0138)
P_9	0.0454***	0.0456***
	(0.0135)	(0.0139)
P_{10}	0.0166	0.0166
	(0.0120)	(0.0114)
P_{11}	0.172***	0.170***
	(0.0232)	(0.0226)
P_{12}	0.0126	0.0114
	(0.0114)	(0.0111)
P_{13}	0.0489***	0.0488***
	(0.0134)	(0.0143)
P_{14}	0.0211**	0.0203**
	(0.00984)	(0.00909)
P_{16}	0.0451***	0.0430***
	(0.0119)	(0.0124)
P_{17}	0.0159**	0.0166*
	(0.00800)	(0.00932)
P_{18}	0.0182**	0.0165**
	(0.00867)	(0.00837)
P_{20}	0.0297***	0.0290***
	(0.0110)	(0.0109)
P_{21}	0.122***	0.104***
	(0.0191)	(0.0167)
P_{22}	0.0731	0.0721
	(0.00370)	(0.00371)
P_{23}		0.00204
		(0.0843)
P_{24}		0.0159**
		(0.00748)
P_{25}		0.0210**
		(0.00894)
P_{26}		0.00516
		(0.00703)
P_{27}		0.00487
		(0.00610)
ϵ	0.0764***	0.0697***
	(0.00336)	(0.00329)
Observations	374	374
Log-likelihood	-2798.3	-2731.3

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$
Standard errors are in parentheses.

APPENDIX B: NORMALIZATION OF THE NOISE PARAMETER (μ) IN ERROR MODELS

The Luce and Fechner models considered in this chapter represent models of "stochastic choice with deterministic preferences", since the true preferences are deterministic but the observed choices are stochastic due to the presence of "white noise" errors (Hey and Carbone 1995). In other words, a particular decision model is assumed to represent true preferences, but errors that cause deviations in observed choices from true preferences are permitted.

In such models, the magnitude of the noise parameter should provide an indication how well the decision model considered can predict observed behaviour (and the extent of the data that is explained by randomness in choice), and it should be comparable across contexts. However, as noted in Section 4.1, the comparison of the noise parameter μ (in its current form) estimates across contexts in which the preference parameters differ has no ordinal or cardinal interpretation. This is because the maximum likelihood estimate of μ depends on the preference parameters included in the model (as well as their estimated values), and so variation in μ does not solely capture differences in noise. Thus, μ estimates cannot be directly compared across decision models within a particular experimental study, or across experimental studies which use different samples (since preference parameter values are expected to vary across samples). Therefore, I propose a normalization procedure which enables the direct comparison of the noise parameter across contexts in which preference parameters differ.

Considering the Fechner error model and following the same notation as in Section 4.1, the latent index used is given by:

$$\nabla EU^{d,q}_{Fech} = \frac{(EU^{d,q}_Y - EU^{d,q}_N)}{\mu}$$

Using this latent index and assuming a standard normal link function, the probability of choosing the Yes insurance option in a particular question is given by:

$$\text{Prob(choose lottery } Y) = \Phi(\nabla EU^{d,q}_{Fech}) = \Phi\left(\frac{EU^{d,q}_Y - EU^{d,q}_N}{\mu}\right)$$

In both equations, EU is replaced by RDU for rank-dependent utility specifications.[111]

Since expected utility (or rank-dependent utility) is a function of preference parameters, its value – as well as the value of $EU^{d,q}_Y - EU^{d,q}_N$ – differs with the values of these preference parameters; as a result, the value of μ which maximizes the log-likelihood function also varies with these parameter values. Further, if different experiments are analyzed, the expected utility differs both because parameter values are likely to differ across samples and because the lottery payoffs and probabilities differ.

[111] See Appendix B in Chapter 1 for a detailed derivation of this model.

The basic issue lies in the fact that the expected utility (assuming a particular decision model and utility function) considered in these structural models carries only an ordinal interpretation and not a cardinal one. However, since the decision problems of the Brazilian experiment (and indeed of most experiments) involve monetary payoffs, the expected utilities also have certainty equivalents, which are expressed in monetary terms. The certainty equivalent is the certain amount of money that is equally desirable to the individual as the risky lottery, given his risk preferences – thus, the individual is indifferent between the two.[112] This provides an appropriate monetary equivalent of a risky lottery and can be compared across contexts – it is often used as a monetary measure of the welfare obtained from, or the attractiveness of, a risky lottery (Cerny 2009). Thus, I utilize this concept in the normalization of the error parameter.

The normalization procedure proposed is as follows. First, the Fechner error model is considered in conjunction with a particular decision model – for simplicity, let us assume the single-parameter EUT-CRRA model.[113] As done before (in Section 4.1), the coefficient of relative risk aversion r and the noise parameter μ are jointly estimated using the structural model (involving the latent index described above) and the maximum likelihood estimation procedure. Given the estimate of r, the expected utility and thus the certainty equivalent of each of the risky lotteries in particular question can be determined – if the utility function used is denoted by U, then for a given EU, $U(CE) = EU$ and thus $CE = U^{-1}(EU)$ (Cerny 2009). Further, since EU is a function of r, CE is also a function of r (Cerny 2009). Thus, for each decision problem d and question question q (of a MPL decision problem in the Brazilian experiment), the certainty equivalents corresponding to both the Yes and No insurance options – $CE_Y^{d,q}$ and $CE_N^{d,q}$ – are calculated. The structural model is then re-estimated, with the expected utilities $EU_Y^{d,q}$ and $EU_N^{d,q}$ replaced by the corresponding certainty equivalent values $CE_Y^{d,q}$ and $CE_N^{d,q}$, and with the noise parameter the only parameter to be estimated. Thus, the latent index is now provided by:

$$\frac{(CE_Y^{d,q} - CE_N^{d,q})}{\mu^*}$$

This latent index is used (along with the standard normal link function) to construct the log-likelihood function, which can then be maximized with respect to the only free parameter, μ^*, to obtain the maximum likelihood estimate of this

[112] Note that the certainty equivalent depends on preferences over risk, that is, on both the decision model assumed as well as the preference parameter values.

[113] This normalization procedure can be easily extended to other decision models and preference functions.

normalized error parameter.[114,115]

The difference between the certainty equivalents of two lotteries in a particular question – as used in the latent index – provides a monetary (and cardinal) measure of the difference in attractiveness between the left and right options, and can be compared across contexts (even if the preference parameter values or lottery payoffs differ). As a result, the estimate of the normalized noise parameter μ^* (which is a latent variable) now provides a comparable estimate of the extent of noise (errors) in decision-making.[116] In this model, the magnitude of μ^* provides an indication of the extent to which choice probabilities depend on the difference in certainty equivalents (with larger values of μ^* indicating lower dependence); in other words, it captures how much variation in the difference between the certainty equivalents of two options (in a particular question) is required to alter an individual's choice. For example, assume that the difference between certainty equivalents (in a particular question) is R$2 and $\mu^* = 0.001$, and this implies very little randomness in choice, that is, the individual will nearly always choose the option with the greater certainty equivalent. However, if $\mu^* = 0.01$, the individual now requires a difference between the certainty equivalents of R$20 to have the same probability (as before) of choosing the lottery option with the greater certainty equivalent. As $\mu^* \rightarrow 0$, the specification collapses to the deterministic model, where the choice in a question is strictly determined by the certainty equivalents of the two lotteries, and the option with the higher certainty equivalent is always chosen.

Certainty equivalents have been used extensively in experimental studies to measure the attractiveness of risky lotteries and the level of risk tolerance (for example, Barr and Packard 2005, Lusk and Coble 2008); Cerny (2009) also use certainty equivalents for normalization in the context of portfolio theory. The idea of using a normalization procedure to compare residuals can be traced back to the Box-Cox transformation technique utilized to normalize the residuals from regression models (for a review of this technique, see Sakia 1992). Additionally, the motivation to obtain a monetary measure of welfare changes has also driven the development of the equivalent and compensating variation concepts, which use the expenditure function to transform changes in utility to monetary terms.

By enabling a comparison of the extent of deviation in observed choices from the predictions of different decision models, this normalization procedure can

[114] Referring to the model outlined in Appendix B of Chapter 1, this latent index implies that choices are made on the basis of *perceived* net certainty equivalent, $CE_Y^{d,q} - CE_N^{d,q} + \varepsilon^*_{d,q}$, where $\varepsilon^*_{d,q}$ represents the "white noise" error, which causes observed choices to deviate from the difference in certainty equivalents. In the case of the standard normal link function, $\varepsilon^*_{d,q}$ is assumed to be normally distributed with mean 0 and standard deviation μ^*, and hence $\varepsilon^*_{d,q}/\mu^*$ has standard normal distribution. Thus, Prob(choose lottery Y) $=$ Prob($CE_Y^{d,q} - CE_N^{d,q} + \varepsilon^*_{d,q} > 0$) $= \Phi\left(\dfrac{CE_Y^{d,q} - CE_N^{d,q}}{\mu^*}\right)$.

[115] The same procedure can be applied to the Luce error model (with Latent Index 2).

[116] This normalization procedure is specific to the case of such structural models of decision-making involving a noise parameter. However, given that numerous studies utilize "white noise" error models and models of "stochastic choice with deterministic preferences" to represent the data-generating process for choices in experimental decision problems, it is important to develop a procedure which enables the noise parameter in these models to be compared across contexts.

be used to determine which decision model best describes the decision-making process under risk of the sample under study. Additionally, it can be used to compare the extent of errors in decision-making across samples. Further, if different elicitation methods are used within the same experiment (as done by Dave et al. 2010, Crosetto and Filippin 2013), this procedure can be used to compare the errors in decision-making associated with the different methods, thus providing an indication of how well each of the methods is understood by participants (which is a crucial determinant of the suitability of an elicitation procedure in a particular setting).

I use this normalization procedure to re-estimate the structural Fechner error specifications for the different decision models using data from the Ethiopian and Brazilian experiments; this procedure utilizes the maximum likelihood estimates of the preference parameters obtained for these specifications in Chapters 1 and 3. Using the Brazilian data, I obtain estimates of μ^* between 3.6 and 3.8 for each of the six decision models analyzed – EUT-CRRA, EUT-EP, PU, RDU-CRRA, RDU-EP and CPT.[117,118] This normalized noise parameter can be compared across the decision models to provide an indication of, for each decision model, the sensitivity of choice probabilities to the differences between the certainty equivalents. Each of these estimates is statistically significantly different from 0, implying a rejection of the hypothesis ($\mu^* = 0$) that there are no stochastic errors in the decision-making process of participants and observed choices are determined purely by the certainty equivalents. However, since the estimated magnitudes of the normalized noise parameter are very similar for all the models considered, there is evidence that each of the models is similarly suitable for describing the true decision-making process of participants, with similar extents of deviations in observed choices from true preferences.

These estimates of μ^* can be compared to those obtained using the Ethiopian data, in order to compare the extent of noise in decision-making across the two experiments. For the estimations of the Fechner error specifications using data from the Ordered Lottery Selection decision problems of the Ethiopian experiment, I use the choice data transformed to binary form, as described in Section 5.2 of Chapter 1 (and as done by Dave et al. 2010). However, as noted in Chapter 1, the Fechner error specification could only be estimated for the EUT-CRRA and EUT-EP models, and thus I only conduct the normalization procedure for these decision models. Additionally, I consider data from the benchmark problem and the insurance problems separately, following the strategy used in Chapter 1. In order to obtain estimates of μ^* that can be compared with corresponding estimates using the Brazilian data, I use historical exchange rate and inflation data to express the certainty equivalents in terms of 2008 Brazilan real, since the monetary payoffs in the Ethiopian experiment are in terms of 2009 Ethiopian birr.

For the benchmark decision problem, the estimates of μ^* are 0.229 (with standard error 0.0103) and 0.271 (with standard error 0.0122) for the EUT-CRRA and EUT-EP models, respectively. Both the estimates are statistically significantly different from 0, indicating a rejection of the hypothesis that there are no

[117] Note that these values differ from the non-normalized μ estimates obtained earlier in this chapter.
[118] Similar results are obtained using the normalization procedure for the Luce error parameter.

stochastic errors in decision-making. However, their small magnitude relative to the μ^* estimates for the Brazilian data indicates that these models provide a relatively better description of the decisions in the benchmark problem of the Ethiopian experiment (than the description of choices in the Brazilian experiment provided by any of the six decision models considered); further, between the two models, EUT-CRRA fits the data better[119], with relatively fewer deviations in observed choices from true preferences (and certainty equivalent differences) – this is in line with the results obtained in Chapter 1.

On the other hand, for the insurance decision problems, the estimates of μ^* are 3.297 (with standard error 0.333) and 0.198 (with standard error 0.00747) for the EUT-CRRA and EUT-EP models, respectively. Once again, the estimates are statistically significantly different from 0. However, the μ^* estimate for the EUT-CRRA model is substantially greater in magnitude than that for the EUT-EP model[120], and is in line with the μ^* estimates obtained using data from the Brazilian experiment. Thus, as found in Chapter 1, EUT-CRRA does not provide a good description of the choices in the insurance problems, with a relatively large fraction of choices attributed to errors in decision-making rather than true preferences; in comparison, EUT-EP provides a significantly better description of choices in the insurance problems. Additionally, EUT-CRRA better describes choices in the benchmark problem than in the insurance problems.

It would be informative to compare the μ^* parameter estimates from the insurance problems of the Ethiopian experiment to those obtained using data from the decision problems of the Brazilian experiment, since both these sets of problems involved index and indemnity insurance purchase decisions. Even though I can only consider the EUT models for the insurance problems of Ethiopian experiment, I find that the μ^* estimates associated with the two EUT models are lower than the μ^* estimates obtained for any of the decision models using the Brazilian data. In particular, the μ^* estimate obtained for the EUT-EP model using the Ethiopian insurance problems is substantially smaller in magnitude than the estimates obtained using the Brazilian data. Thus, even though the four other decision models (with error) could not be estimated for the Ethiopian data, it can be concluded that the decision model (of the six considered) that best fits the insurance problems of the Ethiopian experiment describes the observed choices significantly better (that is, with lesser choices attributed to errors in decision-making) than the model which best fits the Brazilian data can describe choices in the Brazilian experiment.

As a result, there is evidence of significantly greater noise in decision-making in the problems of the Brazilian experiment. Further, if the presence of substantial noise is interpreted as evidence of poor understanding of the decision problems (as indicated by Charness et al. 2013), then the results imply that the simpler Ordered Lottery Selection decision problems of the Ethiopian experiment are better understood than the MPL problems of the Brazilian experiment, by the respective samples. The difference in understanding could be driven by the

[119] Using a Wald test, the hypothesis that the parameter μ^* is equal for the two models is rejected at the 1% significance level.

[120] Using a Wald test, the hypothesis that the parameter μ^* is equal for the two models is rejected at the 1% significance level.

cognitive ability of the subjects or the "cognitive load" placed on participants by the different elicitation procedures (due to their differing complexity). While nearly 71% of participants in the Brazilian experiment were either currently enrolled in, or had completed, college at the time of the experiment, the Ethiopian subjects had, on average, received only four years of formal schooling – given this substantial difference in formal education levels, it is likely that subjects in the Brazilian experiment had, on average, greater exposure to cognitive exercises, and thus greater cognitive ability. Therefore, the results of the comparison of the normalized noise parameter μ^* estimates indicate that the larger errors and poor understanding associated with the Brazilian decision problems are likely due to the greater "cognitive load" placed on subjects by the complex MPL procedure. This is along the lines of the results obtained earlier in this chapter, as well as the observations of Dave et al. (2010) and Crosetto and Filippin (2013).

However, the studies of Dave et al. (2010) and Crosetto and Filippin (2013) compare estimates of the noise parameter μ across contexts in which the preferences parameters differ, without conducting any normalization procedure. Both studies conduct multiple elicitation procedures with the same sample, in order to compare the extent of noise in decision-making for the different procedures through a comparison of the estimated noise parameter (using the Luce and Fechner error models). They find that the estimated preference parameters of utility functions are significantly different across different elicitation techniques[121] – however, they directly compare the estimates of μ obtained for different elicitation techniques in order to determine which technique induces greater noise in decision-making, interpreting a larger magnitude of the estimated μ as representative of greater noise in the decisions elicited using that procedure; further, they attribute the differences in noise to differences in comprehension of the different elicitation techniques. For example, Crosetto and Filippin (2013) find that the estimated r in the EUT-CRRA specification is significantly greater for the Ordered Lottery Selection problem than for the MPL problem in their experiment; they also find that the estimated μ is significantly greater for the MPL decision problem, interpreting this as evidence of greater noise and concluding that this is due to the difficulty of the MPL procedure. However, as highlighted earlier, comparing μ across contexts in which the preference parameters differ could lead to incorrect results and conclusions – therefore, the results of these studies may be misleading, and a normalization procedure is required to enable the comparison of μ.

[121] Charness et al. (2013) also note that elicited risk preferences differ substantially by elicitation procedure.

THE DEMAND FOR INDEX INSURANCE: EXPERIMENTAL EVIDENCE FROM RURAL ETHIOPIA

GAUTAM KALANI AND DANIEL CLARKE*

Abstract

In this chapter, we present a theory outlining the relationship between rational demand for index insurance and wealth. We also analyze data from a framed microinsurance lab experiment involving poor subjects in rural Ethiopia in order to evaluate the validity of this theory. In line with results from recent field studies in developing countries, we find that the demand for index insurance is increasing in wealth at low levels of wealth. However, the relationship between demand and wealth is non-linear, consistent with the hump-shape prediction – outlined in the theory presented – for expected utility maximizers with preferences satisfying risk aversion and constant relative risk aversion. Additionally, we do not find strong evidence that schooling, understanding of the decision problems, or quantitative literacy increase index insurance take-up. Our results indicate that the low take-up of index insurance observed, particularly among the poorest (and most risk averse) individuals, in recent field studies may be a result of rational choice rather than lack of understanding, credit constraints or poor decision-making on the part of rural consumers. We also find no significant association between the take-up of index insurance and background risk.

JEL codes: C91, D01, D03, D81, G22, O16.

*University of Oxford. Emails for correspondence: gautam.kalani@economics.ox.ac.uk and dclarke2@worldbank.org. We would particularly like to thank Marcel Fafchamps, Rocco Macchiavello, Abigail Barr, Sujoy Mukerji and Stefan Dercon for useful suggestions and comments throughout the project. Thanks also to Simon Quinn, Catherine Porter, Erlend Berg, Glenn Harrison, Ruth Vargas Hill, Liam Wren-Lewis, James Martin, Amber Tomas and subjects at seminars at Oxford University. Solomon Ambessie, Zewdu Taddesse Demeke and Solomon Kidane provided invaluable assistance in conducting the experiment. Versions of this paper have been presented at the CEPR Development Economics Symposium 2009 Pre-Conference, a CESS Seminar, a CSAE Seminar and the CSAE Conference 2011. Financial support from the Microinsurance Innovation Facility and the Economic and Social Research Council (ESRC) and logistical support from the Centre for the Study of African Economies (CSAE) and the Ethiopian Development Research Institute (EDRI) are gratefully acknowledged.

1 Introduction

Weather risk is often cited as the most important source of risk in rural regions of developing countries, due to its significant impact on both the level and variability of agricultural income; further, agricultural production is by far the most important source of income for rural households (Cole et al. 2009, Hill et al. 2011). Indeed, Rosenzweig and Binswanger (1993) note that weather risk is the factor contributing to income variability that has the largest impact on consumption and welfare. Additionally, many of the informal risk-sharing strategies utilized (such as inter-household transfers, which occur between geographically proximate households) in these regions are only effective against idiosyncratic shocks, and thus do not facilitate consumption smoothing in the face of covariate weather shocks, which tend to impact all the households in a local geographical area (Giné et al. 2008). In order to be effective against covariate weather shocks, risk pooling needs to occur across a large enough region (for example, across different villages) to ensure imperfect correlation of realizations. Thus, rural households are, in most cases, poorly insured against weather risks such as droughts, floods and even less-severe rainfall shocks (Giné et al. 2008, Hill et al. 2011); additionally, as noted by Dercon et al. (2011), rainfall variability and severe droughts are highly prevalent in rural Ethiopia, and weather risk is a major threat to the welfare of inhabitants.

Given the prevalence of uninsured weather risk and its substantial adverse welfare impact, products that provide insurance against weather shocks are expected to deliver considerable welfare benefits (Hill et al. 2011). As a result, microinsurance provision is expanding rapidly in developing nations and, in particular, innovative products such as index insurance are growing in popularity with policymakers. Over the last decade a variety of institutions have piloted the sale of weather indexed insurance policies to poor farmers, under which the net transfer between insurer and policyholders depends only on readings – that are publicly observable and verifiable – from a contractual weather station;

thus, this net transfer is correlated with, but not a function of, incurred losses. In the case of rainfall indexed insurance, policyholders receive an insurance payout if the rainfall measured at a local weather station falls below a certain level.[1] These policies are expected to help inhabitants of poor rural regions smooth consumption in the face of weather shocks, and thus have the potential to considerably increase household welfare. Additionally, Giné et al. (2008) and Cole et al. (2009) note that index insurance, and the increased transparency it offers, lowers the transaction costs as well as moral hazard and adverse selection problems associated with traditional crop insurance arrangements.

Index insurance has already been piloted in a wide range of developing countries, such as Malawi, Ethiopia, India and Nicaragua (Giné et al. 2008). However, despite the potential risk management and welfare benefits associated with index insurance, voluntary purchase of these products has been much lower than anticipated by proponents. For example, Giné et al. (2008) and Cole et al. (2009), who conduct field experiments in rural India, find that only around 5-10% of participating households purchase index insurance, and a majority of these households purchase enough insurance to cover only 2-5% of agricultural income. Further, these studies find that take-up is lower among the most risk averse individuals as well as the poorest – and therefore most vulnerable – households, who would presumably benefit the most from insurance against weather risk. Similar results have been obtained by Hill et al. (2011), Lybbert et al. (2010) and Giné and Yang (2009) in other developing countries. Indeed, Hill et al. (2011) note that the provision of index insurance has not yet been brought to scale, and the low observed demand for index insurance, particularly among the poorest and most risk averse, has been referred to as a "puzzle" in need of an explanation (Karlan and Morduch 2009).

The focus of this chapter is to analyze the relationship between index insurance purchase and wealth, in order to determine the reason for this low observed

[1] Throughout this chapter, I use the terms weather indexed insurance and rainfall indexed insurance interchangeably.

demand. We present a theory which outlines the relationship between optimal index insurance take-up and wealth, and then empirically test it using data on the choices of participants in a framed microinsurance decision problem of the experiment conducted in Ethiopia (which was also analyzed in Chapters 1 and 2).

The theory presented in this chapter predicts that the optimal demand for index insurance is non-linear, hump-shaped in wealth (as well as risk aversion), for expected utility maximizers with preferences satisfying risk aversion and constant relative risk aversion.[2] This theoretical model specifically takes into account the impact of basis risk (that is, the risk that the net transfer from insurer to policyholder does not match the incurred loss) on the demand for index insurance. Its predictions differ markedly from those of "neoclassical" models of indemnity insurance demand, which predict that insurance take-up should increase monotonically with risk aversion (and decrease with wealth), as well as from those of technology adoption models (such as those described by Feder et al. 1985), which predict that the take-up of new and unfamiliar financial products – such as index insurance – should decrease monotonically with risk aversion (and increase with wealth). These models do not explicitly take into consideration basis risk – which is inherent in index insurance arrangements – but have been proposed as descriptors of index insurance demand by other studies on the subject (for example, Giné et al. 2008, Hill et al. 2011). Indeed, Dercon et al. (2011) find that basis risk, and downside basis risk in particular[3], plays a significant role in the decision to purchase index insurance; therefore, it is crucial to explicitly account for this factor in the development of a theoretical model of

[2] In this chapter, I refer to the wider class of utility functions that satisfy constant relative risk aversion (within the framework of expected utility theory) as expected utility theory constant relative risk aversion (EUT-CRRA) preferences. Thus, the theory and the results in this chapter apply to this wider class of functions (or preferences). This is in contrast to the preceding chapters of this thesis, in which I referred to only a specific constant relative risk aversion functional form as EUT-CRRA utility.

[3] In the case of weather indexed insurance, downside basis risk refers to the possibility that the recorded rainfall is above the predetermined threshold that triggers insurance payouts (and thus the policyholder receives no payout from the insurance arrangement), but the policyholder experiences a bad state of the world (for example, low yield on his plot due to pestilence or because the rainfall on his field is lower than that recorded at the local weather station). Thus, due to downside basis risk, the minimum possible income with the index insurance contract is lower than without it.

index insurance demand.

Using data from the individual index insurance decision problem (T_{IX}), which was framed in terms of a real-world weather indexed insurance purchase decision, we find that the shape of demand for index insurance is consistent with expected utility theory for the class of utility functions satisfying constant relative risk aversion. We find evidence of a hump-shaped relationship between index insurance purchase and wealth, with lower demand from the poorest (who are expected to be the most risk averse) and the richest (who are expected to be the least averse to risk) and highest demand from subjects with intermediate levels of wealth. The finding that demand for index insurance is lower for poorer individuals, and increases with wealth at low levels of wealth, is in line with the empirical findings of Giné et al. (2008) and Cole et al. (2009) in rural India and Hill et al. (2011) in Ethiopia. However, these studies find that the demand for index insurance increases monotonically with wealth and decreases monotonically with risk aversion. They attribute this finding, including the low take-up observed among the most risk averse farmers, to barriers associated with the adoption of a new technology – such as lack of understanding, credit constraints, unwillingness to experiment or poor decision-making on the part of rural consumers – and conclude that the take-up of index insurance can be appropriately described by technology adoption models.

Additionally, unlike Cole et al. (2009) and Hill et al. (2011), we do not find strong evidence that schooling, understanding of the decision problems, or quantitative literacy increase index insurance take-up. Thus, our results indicate that the low take-up observed, particularly among the poorest (and most risk averse) farmers, in these studies may be a result of rational choice – that is, a fundamental lack of desirability of the product due to downside basis risk and actuarially unfair premiums – rather than any of the traditional barriers to the adoption of a new technology. Our results thus contradict the predictions of the "neoclassical" insurance demand model as well as those of technology adoption models, indicating that the theoretical model outlined in this chapter may better inform the take-up of index insurance.

We also find that there is no statistically significant association between the take-up of index insurance and background risk (that is, the risk participants' face in their day-to-day lives outside the experiment), which is not in line with Gollier and Pratt's (1996) risk vulnerability hypothesis. On the other hand, we find some evidence that a household's prior experience with other forms of insurance (such as informal insurance) and the size of its informal risk-sharing network increase the take-up of index insurance.

For index insurance products, the contractual index is not perfectly correlated with the loss – for example, in the case of weather indexed insurance, the rainfall measured at a local weather station is expected to be highly correlated, but not perfectly correlated, with agricultural yields (and thus losses). Index insurance therefore differs from indemnity insurance – unlike indemnity insurance arrangements, index insurance contracts involve basis risk.[4] Thus, for these products, optimal (or rational) demand from an expected utility maximizer is highly sensitive to both the expected claim payment and the claim payment distribution conditional on a large loss having occurred. Therefore, for analyzing the demand for index insurance, field experiments suffer from a critical problem – an objective, unambiguous belief about the joint distribution of losses and index claim payments is extremely difficult to obtain in the field, and neither researcher nor consumer has a precise objective estimate of this joint probability distribution (Clarke 2011).[5] As a result, it is difficult to make clear normative statements about observed demand using field data.

If it is difficult to shed light on index insurance demand through natural field experiments with real index insurance products, we must instead look to environments in which the researcher has a greater degree of control. Laboratory

[4] Indeed, Clarke (2011) notes that index insurance may be better thought of as a derivative or hedging contract.

[5] Even if a researcher had, say, 20 years of matched data for both losses and index claim payments, the required conditional distribution could not be objectively estimated with any degree of accuracy. In practice researchers are likely to have relevant matched data for significantly fewer than 20 years.

experiments represent an important avenue for this purpose – in a lab experiment, such as the Ethiopian experiment analyzed in this chapter, the losses and index are generated by a known randomization device with objective joint probability distribution, and it is therefore possible to make clear normative statements about, and draw conclusions regarding, the demand for index insurance. Additionally, participants in the experiment were chosen from households in the Ethiopian Rural Household Survey (ERHS), which are exposed to substantial weather risk; further, some of these households would be offered real weather indexed insurance policies in the subsequent two years as part of a pilot project.[6]

This chapter therefore contributes to the theoretical and empirical literature on the demand for index insurance in developing countries, and analyzes the correlates of index insurance take-up. It is related to the work of Giné et al. (2008), Cole et al. (2009) and Hill et al. (2011); these studies note that analyzing and correctly establishing the correlates of index insurance take-up is crucial, as it can help shed light on the drivers of index insurance demand and address the causes of low take-up observed in pilot programs. In line with these studies, this chapter focuses on analyzing the shape of index insurance demand and its relationship with wealth, within the framework of expected utility theory.

While acknowledging that index insurance differs significantly from indemnity insurance, these studies do not consider distinct theoretical models of take-up that are specific to index insurance and account for basis risk. Additionally, these studies analyze field and survey data, rather than data from a lab experiment – therefore, they do not possess information on the exact joint probability distribution of the index and losses. Further, they do not test or control for the non-linear effect of wealth on the demand for index insurance, as predicted by the theoretical model in this chapter. In addition, Hill et al. (2011) use data from a module involving a hypothetical insurance contract with hypothetical payoffs contained in the latest round of the ERHS; problems relating to the use

[6] Similarly, subjects involved in the index insurance experiment of Hill and Nobles (2011), which was also conducted in Ethiopia, were offered real index insurance products in the year following the experiment.

of questions involving hypothetical payoffs have been well documented, and risk preferences elicited in non-incentivized tasks may not reflect true risk attitudes, particularly in the domain of financial decision-making (see, for example, Holt and Laury 2002, Charness et al. 2013).

The rest of this chapter is organized as follows. The next section presents the theory outlining the relationship between rational demand for index insurance and wealth (as well as risk aversion). Section 3 discusses the data used in this analysis, while Section 4 considers the factors expected to be associated with index insurance take-up. Section 5 presents the empirical analysis, Section 6 details robustness checks and Section 7 concludes.

2 Theoretical Framework

As is well known, optimal demand for the indemnity insurance is increasing in risk aversion, and under an assumption of decreasing absolute risk aversion, is decreasing in wealth (Pratt 1964).[7] Therefore, optimal insurance take-up in the indemnity insurance decision problem of the Ethiopian experiment (T_{IM}) is, under an assumption of decreasing absolute risk aversion, decreasing in wealth.[8]

However optimal demand for index insurance is fundamentally different to that of indemnity insurance; for expected utility maximizers with preferences satisfying risk aversion and constant relative risk aversion, the optimal demand for index insurance is not monotonic, but rather hump-shaped in risk aversion[9], that is, first increasing and then decreasing in risk aversion (Clarke 2011). Specifically, Theorem 2(i) of Clarke (2011), which applies to both constant absolute relative

[7] As a result, an individual who is infinitely risk averse would rationally purchase full indemnity insurance (Clarke 2011).

[8] Eeckhoudt et al. (1995) note that decreasing absolute risk aversion is a commonly assumed property of preferences, and is fairly unrestrictive, encompassing a wide range of utility functions.

[9] In the context of expected utility theory and constant relative risk aversion considered here, risk aversion refers to the coefficient of relative risk aversion.

risk aversion and constant relative risk aversion[10,11], states that:

> The following conditions on the optimal level of indexed cover hold both for the class of risk averse indirect utility functions that satisfy constant relative risk aversion and for the class of risk averse indirect utility functions that satisfy constant absolute risk aversion, where γ denotes the coefficient of relative or absolute risk aversion, respectively:
> Actuarially unfair products ($m > 1$): $\alpha^*(\gamma) = 0$ for all $\gamma \in (0, \infty)$ if $r \geq p(1 - qm)$. Otherwise $\alpha^*(\gamma)$ is zero for $\gamma \leq \gamma_1$, strictly increasing for $\gamma_1 < \gamma < \gamma_2$ and strictly decreasing for $\gamma_2 < \gamma < \infty$ for some $0 < \gamma_1 < \gamma_2 < \infty$.

In this model, α^* denotes the optimal amount of index insurance cover, p is the probability of the individual loss occurring, q is the probability that the index takes a value which triggers the insurance payout, r is the downside basis risk (that is, the probability that the individual will incur a loss but the index value will not trigger a payout), and the index insurance product is priced with a multiple[12] of m.

The hump-shaped relationship between optimal purchase and risk aversion is caused by the combination of actuarially unfair premiums, whereby premiums are greater than expected claim income, and basis risk, the risk that the income from index insurance will not accurately reflect the incurred loss. These factors cause demand to be low from both the risk neutral, for whom insurance purchase decreases mean income, and the very risk averse, for whom index insurance purchase decreases the minimum possible income (due to downside basis risk). It is only optimal for those with intermediate levels of risk aversion to purchase

[10] The theoretical model developed by Clarke (2011) is structurally similar to Doherty and Schlesinger's (1990) model of indemnity insurance contracting with a non-zero probability of insurer default (that is, a risk of contractual non-performance).

[11] In this chapter, we focus on the predictions of this theorem with respect to expected utility theory and constant relative risk aversion preferences, a specific functional form of which has been analyzed in the preceding three chapters of this thesis as well.

[12] This multiple is defined as (Premium Charged)$/$(Expected Claim Income).

index insurance.[13,14]

This intuition also follows through to the relationship between optimal demand for index insurance and wealth, so long as preferences satisfy decreasing absolute risk aversion and risk aversion is sufficiently high at low levels of wealth and sufficiently low at high levels of wealth.[15,16] As proved in Theorem 1 of Clarke (2011), for sufficiently high risk aversion (low wealth) optimal demand will be zero, as index insurance purchase decreases the minimum possible income. Similarly for preferences sufficiently close to risk neutral (high wealth), demand

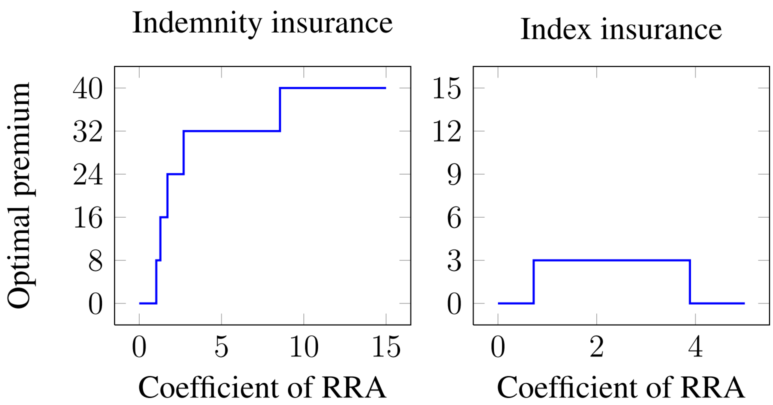
[13] Note that this theory only applies to risk averse individuals. However, in the context of poor inhabitants of rural Ethiopia and decisions involving significant amounts of money (mean realized earnings from the Ethiopian experiment was two to three days' income from casual farm labour), risk loving behaviour has little conceptual justification. Additionally, the results in Chapters 1 and 2 indicate that subjects in the Ethiopian experimental sample exhibit moderate risk aversion in the decision problems. Indeed, Charness and Viceisza (2012) note that risk loving behaviour has rarely been observed in experimental studies, and when it is observed, is indicative of a lack of understanding of the decision problems on the part of subjects (leading to "frivolous" responses), rather than reflective of true risk preferences. Thus, while this theory applies only to risk averse individuals, it is highly likely that the entire sample – comprising of experiment participants who are relatively poor – falls within this category.

[14] Thus, the results (predictions) of this theoretical model differ significantly from those of the "neoclassical" model pertaining to indemnity insurance demand (for example, the model described by Giné et al. 2008). Table 5 in Clarke (2011) summarizes the key results of the "neoclassical" indemnity insurance model and the index insurance model – the table shows that the two models produce very different predictions. Most importantly, the indemnity insurance model predicts a monotonically increasing relationship between insurance demand and risk aversion, while the index insurance model of Clarke (2011) predicts a hump-shaped non-monotonic relationship. Further, the reason for this difference in results is basis risk – while the "neoclassical" indemnity insurance model does not incorporate it, the index insurance model does. This is because indemnity insurance does not involve basis risk, since in an indemnity insurance arrangement, policyholders receive an insurance payment when they face a bad individual outcome and thus income from insurance accurately reflects the incurred loss. On the other hand, basis risk is inherent in any index insurance arrangement, and mathematically, the treatment of basis risk in the index insurance model of Clarke (2011) is similar to the treatment of contractual non-performance in Doherty and Schlesinger's (1990) model of indemnity insurance contracting with a non-zero probability of insurer default.

[15] Note that constant relative risk aversion (CRRA) is a special case of decreasing absolute risk aversion, and CRRA represents a useful parametrization of a set of decreasing absolute risk aversion functions, ordered in terms of risk aversion.

[16] In Chapter 2, analyzing data from the Ethiopian experiment, I do find evidence that wealth has a significant negative effect on risk aversion.

Figure 1. *Optimal take-up for T_{IM} and T_{IX} under constant relative risk aversion*



for actuarially unfair insurance will also be zero.[17,18] It is only optimal for those with intermediate levels of wealth to purchase index insurance.[19]

This intuition follows through to our specific decision problems. As can be seen in Figure 1, optimal take-up for the indemnity insurance problem (T_{IM}) under constant relative risk aversion is increasing in risk aversion, but optimal take-up for the index insurance problem (T_{IX}) is first increasing, then decreasing in risk aversion.[20,21]

To test these theoretical predictions we include a measure of wealth and its square as correlates of index insurance take-up in T_{IX}, that is, as explanatory variables

[17] For a risk neutral decision-maker, expected wealth (and therefore expected utility) will be strictly higher on purchase of zero units of actuarially unfair index insurance than on purchase of any strictly positive number of units, and this will also be true for sufficiently smooth preferences sufficiently close to risk neutral. For example, note that Theorem 2(i) of Clarke (2011) proves that optimal demand for actuarially unfair index insurance will be zero for constant absolute or constant relative risk aversion if the coefficient of constant absolute or constant relative risk aversion is sufficiently close to zero.

[18] With zero demand for risk neutral (high wealth) and infinitely risk averse (low wealth) individuals, demand for index insurance cannot be monotone unless it is zero everywhere.

[19] Additionally, Theorem 4 of Clarke (2011) states that for any individual with preferences satisfying risk aversion and decreasing absolute risk aversion, there is an upper bound to the optimal level of insurance purchase – this is because if the individual is averse enough to risk to want to purchase the cover, he would also be averse enough to the downside basis risk (that is, the risk that the net transfer from insurer to policyholder does not match the incurred loss) to limit the size of the hedge.

[20] See Tables 1 and 2 in Chapter 1 for the outcome payoffs and associated probabilities in these decision problems.

[21] The discrete jumps in Figure 1 arise because there were only six discrete options for the amount of insurance purchase in the decision problems, and the value of the coefficient of relative risk aversion (γ) corresponding to such a jump is that at which the expected utilities provided by the two options are equal (that is, individuals with that level of risk aversion are indifferent between the two amounts of insurance purchase).

in specifications with the respondent choice in T_{IX} as the dependent variable.[22] If the coefficient estimate on the level term is significantly positive and that on the squared term is significantly negative, we can reject the hypothesis that index insurance take-up in T_{IX} is monotonically decreasing in wealth; further, if the coefficient estimates of the level and squared wealth terms imply a maximum take-up of index insurance at a point near the mean of the wealth distribution, this would provide evidence of a hump-shaped relationship between take-up and wealth (as opposed to just a concave relationship).

3 Data

This chapter, like Chapter 2, uses combined data from the Ethiopian lab experiment and the ERHS. The experiment included a benchmark decision problem, framed in the abstract, and four framed insurance decision problems. The focus of this chapter is to analyze the take-up of index insurance. Hence, only one of the four insurance decision problems – the problem T_{IX} which involved the purchase of index insurance – is analyzed in this chapter. 378 subjects participated in the experiment – while all 378 subjects were involved in the benchmark decision problem, only 258 were involved in the individual index

[22] As mentioned earlier, the Clarke (2011) model also predicts that index insurance take-up is hump-shaped in risk aversion, and the impact of wealth on take-up in this model is driven purely by risk aversion. Though we can estimate the coefficient of relative risk aversion γ from participants' choices in the benchmark decision problem of the Ethiopian experiment (B), we cannot directly test for the existence of a hump-shaped relationship between index insurance take-up and γ. This is because each choice in the benchmark problem corresponds to a range of implied γ, and the only way to reduce the ranges to point values is to use a rather arbitrary averaging technique, as suggested by Binswanger (1981). In particular, the averaging technique cannot appropriately account for ranges of γ that are right- or left-unbounded, such as those corresponding to choices A and F in the benchmark decision problem. Binswanger (1981), for example, considers the corresponding point value to be 12% above (below) the lower (upper) bound for the right (left) unbounded range, though there is no theoretical justification for this. In addition, this discrete, arbitrarily chosen point value of γ can only take on very few (six) values, and including the level and squared terms of this discrete variable to test the non-linear relationship would not be very informative or appropriate. Therefore, we prefer to test the prediction of the Clarke (2011) model that the relationship between index insurance take-up and wealth is hump-shaped, as including the level and squared terms of this continuous variable would be more informative. Further, given that we only expect the impact of wealth on take-up in the experimental setting to be through risk aversion (as explained in the next section), testing the relationship between index insurance take-up and wealth can provide strong implications for the relationship between take-up and risk aversion.

insurance decision problem (T_{IX}). T_{IX} was framed to be as similar as possible to a real weather indexed insurance purchase decision, and participants were given an explanation for the probability structure in a real-world agricultural context. In the index insurance problem, which was framed in the loss domain, subjects chose one out of six options, each corresponding to a different amount of index insurance purchase and premium payment.[23] Further, T_{IX} involved actuarially unfair premiums, which is the norm for index insurance contracts (for example, those offered in the pilot projects of Giné et al. 2008, Cole et al. 2009).[24,25] A complete description of the individual index insurance decision problem, as well as the other decision problems of the experiment, is provided in Section 4 of Chapter 1. Further, Table 2 in Chapter 1 details the outcome payoffs and associated probabilities in T_{IX}, while Figure 5 presents the visual aid used for this decision problem.

As mentioned in Chapter 1, whilst experimental economists might argue that the compound probability structure for the index insurance treatment may be too complex for subjects to understand, we consider it to be much less difficult to understand than the joint probability structure for a real weather indexed insurance policy; although farmers might have a good understanding of the marginal loss distribution for their farms, they are unlikely to have a good understanding of the conditional distribution of weather indexed claim payments (Giné et al. 2007, Hill and Nobles 2011).

[23] In T_{IX}, choice A involved the least amount of insurance purchase and choice F involved the most insurance purchase.

[24] The index insurance in T_{IX} was priced with a loading of 20%, where an insurance loading is defined as (Premium Charged)/(Expected Claim Income) $-$ 1. This loading is low compared to reported commercial loadings for crop insurance, which range from 70% to 430% for weather indexed insurance (Cole et al. 2009, Table 1) and 140% to 470% for indemnity insurance (Hazell 1992, Table 1). However, the 50% probability of claim payment in the experiment is much higher than that for commercial insurance products, and so these loadings cannot be directly compared.

[25] In the case of T_{IX}, the downside basis risk is the probability that a blue bag (good index) is selected – so there is no insurance payout – but a yellow token (bad individual outcome) is drawn from the blue bag.

3.1 Experimental sample and matched ERHS data

The subjects in the experiment, which was conducted in November and December 2009, were chosen from seven sites of the Ethiopian Rural Household Survey (ERHS), spanning three regions of the country. In addition to subjects' choices in the various microinsurance decision problems, the experiment dataset contains information on literacy, schooling, quantitative literacy, understanding of the decision problems, occupation as well as various other demographic characteristics of the subjects. The ERHS dataset consists of seven rounds of detailed longitudinal survey data on various socioeconomic and demographic characteristics of 1,350 households, such as asset ownership, income, consumption, membership in risk-sharing groups and household size. For the purpose of this chapter, we use data mostly from the latest (seventh) round of the survey, conducted between April and August 2009, since the timing closely matches that of the experiment. However, we also utilize certain panel aspects of the data, as will be described in the next section.

Data from the ERHS was matched with data from the experiment using unique household identification numbers. This combined dataset provides information on an extensive set of factors which are expected to be associated with index insurance take-up in the experiment, and thus enables a detailed analysis of the correlates of take-up. These potential correlates of index insurance demand are described in the following section.

4 CORRELATES OF INDEX INSURANCE TAKE-UP

Wealth affects index insurance purchase in two important ways. First, a household's wealth is indicative of the credit and liquidity constraints it faces, and these financial constraints may play a key role in its decision to purchase insurance in the field; Giné et al. (2008) and Cole et al. (2009) find that index insurance take-up increases monotonically with wealth, and conclude that this is because wealthier households have lesser constraints on credit and liquidity. Second, household wealth is an important determinant of risk aversion, which in turn affects take-up of insurance. In an experimental setting where each subject is given 65 birr and the option to purchase insurance to offset adverse outcomes, the credit (or liquidity) constraint effect of wealth is not expected to impact insurance take-up (as it may do in the field), and wealth would only affect the purchase of index insurance through risk aversion. As detailed in Section 2, if subjects' preferences over aggregate wealth satisfy decreasing absolute risk aversion, then risk aversion decreases with wealth and the hump shape of demand relative to risk aversion is expected to carry over to wealth; both poor subjects (those with high risk aversion) and rich subjects (those with low risk aversion) would have low demand for index insurance due to basis risk and actuarially unfair premiums respectively, leaving higher demand only for those with intermediate levels of wealth.[26] To test the theoretical prediction that the demand for index insurance is hump-shaped in wealth, we include a measure of wealth and its square as correlates of index insurance take-up in T_{IX}. Additionally, given that the laboratory setting enables isolation from the credit and liquidity constraint effects of wealth, and we only expect the impact of wealth on take-up to be through risk aversion, this would also provide strong implications for the relationship between take-up and risk aversion.

The hump-shaped prediction for optimal demand differs significantly from predictions of the "neoclassical" mean variance model of Giné et al. (2008), which does not account for basis risk and under which purchase of index insurance is expected to be monotonically increasing in risk aversion (and therefore decreasing in wealth); within the framework of this model, the most risk averse (and the poorest) would be expected to benefit the most from the consumption smoothing provided by insurance, and thus would be expected to

[26] Hill et al. (2011) also note that in this region, the most risk averse individuals are likely to be the poorest. Further, the results in Chapter 2 – which analyzes data from the Ethiopian experiment and ERHS – indicate that wealth has a negative impact on risk aversion, and are in line with the results of recent experimental studies conducted in developing countries (for example, Yesuf and Bluffstone 2009, Tanaka et al. 2010, Tanaka and Munro 2012).

have the strongest demand for any insurance product. The prediction of the hump-shaped relationship between index insurance demand and wealth is also different from the prediction provided by the literature on technology adoption, according to which the take-up of a new technology is expected to be monotonically decreasing in risk aversion and increasing in wealth (a comprehensive review of this literature is provided by Feder et al. 1985). This is primarily due to the aversion to uncertainty associated with the transaction and benefits (which may be stochastic and dynamic) of the new, unfamiliar technology, combined with uncertainty in potential buyers' perceptions of these benefits (Lybbert et al. 2010, Hill et al. 2011). Giné et al. (2008) claim that this argument also applies to new and unfamiliar financial products, such as index insurance (which ERHS households had not been previously offered). Additionally, Hill et al. (2011), analyzing data from Ethiopia, conclude that models of technology adoption can inform the purchase of weather indexed insurance.

For a particular decision problem, all subjects face the same risky outcomes and are given the same choices. However, subjects differ significantly in the risk they are exposed to in the natural course of their lives outside the experiment (labelled "background" risk), and subjects' earnings in the experimental decision problems are statistically independent of this background risk. Gollier and Pratt (1996) derive conditions, which they label "risk vulnerability", under which the addition of an independent background risk would cause an increase in aversion to other risks (that is, an increase in indirect risk aversion); further, they note that constant relative risk aversion (CRRA) preferences are risk vulnerable. In this chapter, we are particularly interested in decision-makers with utility functions that satisfy CRRA, and thus Gollier and Pratt's (1996) risk vulnerability hypothesis predicts that greater background risk should be associated with lower risk-taking in the index insurance decision problem. On the other hand, Diamond (1984) and Quiggin (2003) investigate conditions under which aversion to a particular risk would be reduced by the addition of an independent background risk. Indeed, Lusk and Coble (2008) note that determining the sign and magnitude of the relationship, if any, between background risk and risk aversion – and therefore insurance purchase decisions which depend intimately on risk aversion – is ultimately an empirical issue, as theoretical expositions do not provide a clear prediction of this association. However, there are relatively few empirical studies on the subject, and results from these studies are inconclusive (Lusk and Coble 2008). Thus, Tanaka and Munro (2012) stress the importance of further experimental research analyzing the association between background risk and risk attitudes as well as experimental decision-making.

Therefore, to investigate whether background risk is related to decision-making in the index insurance problem of the Ethiopian experiment, we include a measure of the background risk faced by the participant as a correlate of respondent choice in T_{IX}. Hill et al. (2011) also include a measure of the consumption risk faced by participants in their day-to-day lives as a correlate of take-up of their hypothetical index insurance instrument.

Further, following Giné et al. (2008), Cole et al. (2009) and Hill et al. (2011), we also include controls for membership in social groups and access to risk-sharing networks, cognitive ability and understanding of the decision problems, and demographic characteristics of the participants (such as age and gender). These variables, and the motivation for including them in the analysis of index insurance take-up, are described in detail in the next section.

For the index insurance decision problem, we apply both an Ordinary Least Squares (OLS) model and an Ordered Probit model to the data.[27,28] The dependent variable in these specifications is the choice of the subject in problem T_{IX}, and thus can take on values 1, 2, 3, 4, 5 or 6, with greater values indicating greater index insurance purchase, that is, greater premium choice (refer to Table 2 in Chapter 1).[29] The Ordered Probit model is included since the dependent

[27] The Ordered Probit model estimates coefficients using a maximum likelihood estimator – see Wooldridge (2002) for more details on the model.

[28] Giné et al. (2008) and Hill et al. (2011) only possess data on whether or not index insurance was purchased by an individual, rather than the amount of cover purchased; as a result, they can only estimate Probit specifications.

[29] The choices A through F in decision problem T_{IX} are labelled as choices 1 through 6.

variable is an ordinal, ordered response that is not continuous and so the linear structure of the OLS model may lead to an unbounded range for the coefficient estimates, even though the dependent variable can only take on values 1, 2, 3, 4, 5 or 6 (Wooldridge 2002).

4.1 Description of variables

The measure of household wealth used in this analysis is tropical livestock units (TLUs).[30] TLUs are standardized units of different types of livestock, and they are used as a measure of total livestock ownership in numerous studies set in the context of developing economies (for example, Dercon 2004, Barrett and McPeak 2006).[31] Dercon (2004), in a study of growth and shocks in rural Ethiopia, observes that livestock typically accounts for over 90% of the value of household assets and is the most marketable asset in this region – therefore, in rural Ethiopia, as in many of the poorest rural regions of the world, livestock ownership is an appropriate and important measure of household wealth, while tropical livestock units provide a suitable and comparable description of livestock ownership. Table 1 provides the summary statistics for livestock ownership as well as the other ERHS and experiment variables used as explanatory variables in the analysis of the correlates of index insurance take-up in rural Ethiopia. The summary statistics are presented separately for the full experiment sample and the sub-sample involved in decision problem T_{IX}.

The households of the 378 participants in the experiment own 10.5 TLUs each, on average, while those of the 258 subjects involved in T_{IX} own 10.8 TLUs each, on average. Both these values are slightly greater than the mean for the complete ERHS round 7 sample, and all households involved in the experiment own at least some livestock; TLUs for households in our experimental sample

[30] This is one of the measures of household wealth used in Chapter 2, and is also used by Lybbert et al. (2010) in their analysis of index insurance take-up in an experiment conducted in Northern Kenya.

[31] TLUs provide a single figure that expresses the total amount of livestock owned in a common unit. For the purposes of the ERHS, it is calculated using the following conversions: oxen=1, cows=0.70, bulls=0.75, horses=0.50, goats=0.10, sheep=0.10 and other similar values (Dercon 2004).

Table 1. *Summary statistics*

No. of Participants	T_{IX} (258)		Full Experiment Sample (378)	
	Mean	Std. Dev.	Mean	Std. Dev.
ERHS Variables				
Tropical livestock units	10.81	2.788	10.54	2.900
Std. dev. of consumption[a]	546.2	329.1	528.8	335.4
No. of iddir	2.128	1.555	2.246	1.643
Can obtain 100 birr in emergency	0.829	0.377	0.827	0.379
If equb member	0.272	0.446	0.247	0.432
Household size	5.630	2.245	5.551	2.341
Experiment Variables				
Understanding	0.861	0.173	0.870	0.172
How many decision problems paid for?	0.624	0.485	0.653	0.477
Which decision problem paid for?	0.926	0.262	0.934	0.249
Which colour token is bad?	0.977	0.151	0.976	0.153
Wheel spin or bag draw first?	0.919	0.274	0.894	0.308
No. of yellow tokens in yellow bag?	0.868	0.339	0.880	0.324
Yellow or blue token draw more likely?	0.868	0.339	0.884	0.321
Quantitative literacy	0.495	0.205	0.514	0.210
$5+3=?$	0.822	0.384	0.862	0.345
$3\times 7=?$	0.523	0.500	0.545	0.498
$\frac{1}{10}$th of $300=?$	0.252	0.435	0.300	0.458
5% of $200=?$	0.0116	0.107	0.0132	0.114
Riskier to plant one crop or multiple crops?	0.868	0.339	0.852	0.356
If literate	0.790	0.408	0.770	0.421
Schooling obtained (years)	4.702	3.892	4.149	3.795
Age (years)	46.60	15.86	45.14	15.94
If female	0.322	0.468	0.325	0.469
If household head	0.702	0.458	0.696	0.461
If farmer	0.682	0.467	0.664	0.473
Fraction of earnings kept	0.468	0.462	0.432	0.463

[a] Measured in 1994 Ethiopian birr

range from 8 to 31.5, as compared to 7 to 31.5 for the entire ERHS sample.[32] In Section 6, we show that our results are also robust to an alternative measure of wealth, calculated as the first principal component from a principal component analysis of asset ownership variables in the ERHS survey.[33]

Exploiting the panel aspect of the ERHS data, we use the standard deviation of household consumption over all seven rounds of the ERHS (conducted between 1994-2009) as a measure of the level of background risk to consumption faced by a participant's household in the recent past. Numerous studies use measures based on the inter-temporal standard deviation of consumption or income as measures of risk and shocks affecting rural households (for example, Jalan and Ravallion 1998, Kamanou and Morduch 2004). In this case, the standard deviation of consumption is used rather than that of income because the ERHS income data is far less reliable than the consumption data (Porter 2011). In addition, we want to measure the final level of risk faced by the household – that which could have a significant impact on household welfare – after it has used all available insurance and consumption smoothing measures (including possibly-inefficient asset investment and divestment), which would not be captured by the standard deviation of income.

In the context of decisions in the experiment, this measure can be considered as background risk, as it is uncorrelated with the earnings and "foreground" risk in the index insurance decision problem and is difficult to insure against (since we use consumption, which provides a measure of welfare after the use of available insurance mechanisms and risk-coping strategies to smooth income). Guiso and Paiella (2008) and Hill et al. (2011) also use variables based on the inter-temporal deviation of household consumption (with the latter using data from the rounds of the ERHS between 1994-2009) to measure the background

[32] Table 3 in Chapter 2 provides the summary statistics for the complete ERHS round 7 sample.

[33] Land ownership is another commonly-used proxy for wealth in studies using the ERHS dataset (for example, Porter 2011, Hill et al. 2011), and is used in Chapter 2 as well. When we use the total land owned by the household (measured in hectares), instead of TLUs, as the measure of wealth in the analysis conducted in Section 5, there are no substantive differences in the results.

risk faced by individuals in their samples.[34]

Consumption is measured as the total monthly household consumption in 1994 Ethiopian birr. It includes the consumption of food, purchased food and non-investment non-food items (that is, it excludes expenditure on durables, health and education) – as noted by Porter (2011), this measure has been utilized by various other studies of consumption and poverty conducted using the ERHS dataset. Table 1 shows that the inter-temporal standard deviation of monthly consumption is relatively high for households participating in the experiment – for these households, the mean standard deviation of around 530 birr represents a significant fraction (just under half) of the average monthly consumption measured in round 7 of the ERHS (approximately 1100 birr). This indicates that subjects in the experiment face considerable risk to consumption, just like most households included in the survey.

Giné et al. (2008) and Hill et al. (2011) claim that the literature on technology adoption can provide insights into the purchase and take-up of new and unfamiliar financial products, such as weather indexed insurance.[35] In accordance with models of technology adoption, Giné et al. (2008) hypothesize that when a new technology or financial product is introduced to rural farmers, households draw inferences based on experience and familiarity with other similar products. In particular, Lybbert et al. (2010) and Hill et al. (2011) note that prior experience with financial products would increase trust in the financial sector (as well as the perceived probability of receiving payment from the product) and improve understanding of a new product such as index insurance, thus facilitating its take-up.

In addition, the literature on technology adoption and diffusion indicates that households rely heavily on the large information flows between members of

[34] However, it is important to note that while the inter-temporal deviation in household consumption does capture background risk faced by the participant, all of the deviation in consumption is not due to risk – for example, some of the inter-temporal variation in household consumption could be caused by life cycle effects or household decisions regarding the purchase of durables.

[35] Feder et al. (1985) provides a comprehensive survey of this literature, including the theoretical models of technology adoption and empirical evidence on the subject.

social groups (or networks) in deciding whether to take up a new technology or financial product (Feder et al. 1985, Bandiera and Rasul 2006); thus we would expect individuals with larger social or informal risk-sharing networks to have more experience with the adoption of new products, making them more likely to purchase index insurance. Dercon et al. (2011) note that while trust (as well as belief in the credibility of plans of the insurance vendor) and understanding are important determinants of insurance purchase decisions in poor rural economies, such groups overcome some of the barriers associated with the take-up of new products such as index insurance, and therefore members of these groups are more likely to adopt new products than non-members. Therefore, in line with the work of Giné et al. (2008) and Hill et al. (2011), we hypothesize that farmers with larger social networks and greater experience with financial products are more likely to purchase index insurance in the experiment.[36]

We include ERHS variables to capture the subject's membership in social groups, prior experience with financial products (such as savings and insurance), and access to informal insurance networks. These include dummy variables indicating whether the household was part of an equb group (a mutual savings association, similar to a ROSCA) and whether the household could obtain 100 birr (8 USD) within a week in case of an emergency (which captures the household's access to informal insurance, and therefore is indicative of the size of its informal risk-sharing network). Further, the number of iddir groups that the subject's household was a member of is used as an indicator of both the size of the household's social network as well as the household's, and thus the subject's, prior experience with informal insurance. Iddirs are informal insurance groups indigenous to Ethiopia that were originally formed to cope with the relatively high cost of funerals – however, currently many iddirs also provide informal insurance

[36] On the other hand, if we consider index insurance to be one of many risk management strategies (that is, an element of a risk management portfolio) used by inhabitants of rural Ethiopia and if index insurance is considered by potential buyers as a substitute for other strategies, we may expect those individuals with greater access to relatively cheap informal and formal risk-coping options to have lower demand for index insurance (Hill et al. 2011). However, this is not expected to be an important factor in the experimental data analyzed in this chapter, since the index insurance contract is the only option available to participants for mitigating risk in the experimental decision problem.

and credit to members when they experience other adverse shocks, such as fires, illnesses or loss of livestock (Hoddinott et al. 2005).[37] Hoddinott et al. (2005) observe that iddir members have larger social networks, and also better access to informal insurance, than non-iddir members.

Access to formal financial products – such as savings, credit and insurance – in rural Ethiopia is severely limited; most inhabitants do not have bank accounts or experience with formal financial institutions (Dercon 2004).[38] Thus, in order to proxy for experience and familiarity with financial transactions, I use the above-mentioned variables which capture experience with informal savings and insurance, as well as access to social groups and informal risk-sharing networks. Hill et al. (2011) also use similar variables to capture the availability of informal insurance options and the size of social networks, noting that these channels encompass the main *ex post* risk-coping strategies available to rural Ethiopians.[39]

As noted by Dercon et al. (2011), iddirs are widespread in rural Ethiopia, with almost every household a member of an iddir. Indeed, nearly 96% of the households of participants in the experiment were members of at least one iddir. Further, Table 1 shows that the households of experiment participants were members of two iddirs on average; on the other hand, only 25% of these households were members of an equb group.[40] Additionally, a large fraction (nearly 83%) of the households of experiment participants reported that they could obtain 100 birr within a week in case of an emergency.

[37] The functioning of iddirs is independent of any government or NGO involvement (Dercon et al. 2011).

[38] For example, only 6% of households in the ERHS sample hold savings in a bank account (Hill et al. 2011).

[39] Further, it is important to note that while iddirs represent an important source of informal insurance in rural Ethiopia, assistance against risk can also come from individuals who are not members of the same iddir – for example, relatives, neighbours and fellow members of labour-sharing groups; thus, while there may be some overlap, the 100 birr dummy variable and iddir (as well as equb) membership variable do not necessarily capture the same source of insurance.

[40] Equbs are not as widespread as iddirs in rural Ethiopia, as the summary statistics indicate. Most households that were equb members were members of only one equb. Therefore, the number of equbs that a household was a member of was not recorded in the ERHS, and we only use a dummy variable indicating whether the household was a member of an equb or not. In contrast, iddir membership in the region is extensive and many households involved in the survey were members of multiple iddirs (Hoddinott et al. 2005). Thus, to distinguish between households on the basis of iddir membership, we use the number of iddirs that the household was a member of (which was recorded in the ERHS) rather than simply a dummy indicating whether the household was a member of at least one iddir.

Giné et al. (2008) and Cole et al. (2009) argue that cognitive ability, quantitative literacy and understanding of financial products such as index insurance are instrumental in their proper valuation and rapid take-up. They argue that with the introduction of a new product which may not be properly understood – like index insurance – individuals with "greater cognitive ability to understand the product are more willing to experiment with it". Similarly, Hill et al. (2011) report that a number of studies on technology adoption have found that educated individuals are more likely to be early adopters of new technologies. Lybbert et al. (2010) observe that index insurance products are relatively complex, with benefits that are stochastic, while Dercon et al. (2011) note that the presence of basis risk makes the decision of whether or not to purchase index insurance much more complicated. Thus, comprehension of index insurance is a potential factor in determining its take-up.[41]

In conducting the experiment, enumerators went to great lengths to make all decision problems understandable to both literate and illiterate subjects, spending a substantial amount of time providing detailed explanations of each decision problem to subjects, as well as using visual aids and tangible randomization devices. However, in spite of this, it is quite likely that people with different education levels and literacy status had different perceptions of the decision problems. Therefore, we include as explanatory variables both the years of schooling obtained by the experimental subjects and a literacy dummy, to test whether greater cognitive ability is associated with more index insurance take-up.[42]

While 77% of the experimental subjects were literate, the average subject had only around four years of formal schooling. Though Lybbert et al. (2010) note that comprehension is expected to be highly correlated with education, it

[41] Additionally, Lybbert et al. (2010) note that such experiments conducted in the field could build understanding and help individuals in the region learn about relatively complex new products such as index insurance, possibly leading to greater take-up of the product on its introduction.

[42] At the end of the experiment we asked each subject to sign a receipt for their earnings and they were offered both a pen and a fingerprint card. The literacy dummy variable is one if the subject signed with a pen and zero if with a thumb print. Enumerators did not prompt.

is possible that years of formal schooling is a poor proxy for the ability to solve the mathematical problems that rural inhabitants encounter in everyday life, and for whether individuals possess a good intuitive understanding of how insurance might help them (Cole et al. 2009); further, Dave et al. (2010) stress the importance of distinguishing between mathematical skills and general intelligence as the two are likely to have distinct effects. Therefore, in line with the work of Cole et al. (2009) and Lybbert et al. (2010), we also include direct measures of understanding and quantitative literacy elicited in the experiment – these are the fraction of six questions relating to the understanding of the decision problems answered correctly and the fraction of five questions assessing probability and mathematical skills answered correctly (refer to Table 1 for more details on these questions). Dave et al. (2010) note that such measures of quantitative literacy appropriately capture mathematical skills and the proficiency with numbers in everyday life. Further, the quantitative literacy measure is indicative of participants' understanding of insurance, which is expected to affect the take-up of index insurance (Giné et al. 2008) – failure on the quantitative literacy questions suggests a poor understanding of insurance products and their working. While the subjects exhibited a fairly good understanding of the decision problems (correctly answering 87% of the understanding questions, on average), each subject answered only 51% of the quantitative literacy questions correctly, on average.[43] Though this measure indicates a relatively low level of quantitative literacy, it is still higher than that measured by Cole et al. (2009), who find that the average respondent in their sample of rural inhabitants of Gujarat (India) correctly answered only 34% of quantitative literacy questions of a similar difficulty.

[43] The final quantitative literacy question listed in Table 1 asks whether it is riskier for farmers to plant one crop or multiple crops. While it is generally expected that diversification reduces risk exposure (and hence planting multiple crops is considered as the correct answer to this question in the formulation of the quantitative literacy variable), in certain cases farmers in poor rural economies can improve risk coping through specialization, that is, by planting a single crop which is resistant to pests, droughts and other environmental risk factors (Fafchamps 2003). For example, Fafchamps (2003) notes that in many areas of West Africa, millet is the only cultivated crop, and specialization in the cultivation of this single, robust crop is less risky and forms the main source of income smoothing. Therefore, the correct answer to this particular quantitative literacy question is not immediately clear. However, the results for all the specifications in this chapter remain substantively the same when the quantitative literacy measure excludes this question.

Other characteristics of the subjects that are expected to affect index insurance take-up are also considered – these include gender (represented by the dummy variable 'If female' taking the value one if the subject was female), whether the subject's primary occupation was farming or not (indicated by the dummy variable 'If farming', which takes the value one if the subject was a farmer), whether the subject was a household head or not (indicated by the dummy variable 'If household head', which takes the value one if the subject was a household head), and household size.[44] These are similar to the demographic characteristics included by Giné et al. (2008), Cole et al. (2009) and Hill et al. (2011) in their analyses of the correlates of index insurance take-up.

In addition to these variables, the self-reported fraction of the total earnings from the experiment that the subject intended to keep for himself is also included as a correlate of index insurance take-up. We hypothesize that those subjects keeping a larger fraction for themselves are expected to choose riskier options in the experiment than subjects who intended to share their experimental earnings with others, and hence were wary about returning with very little (or no) money from the experiment other than the participation fee of 5 birr. Along similar lines, Hill et al. (2011) include a self-efficacy measure, based on survey participants' reported responses on whether they felt they had the power to make decisions that could change the course of their lives.

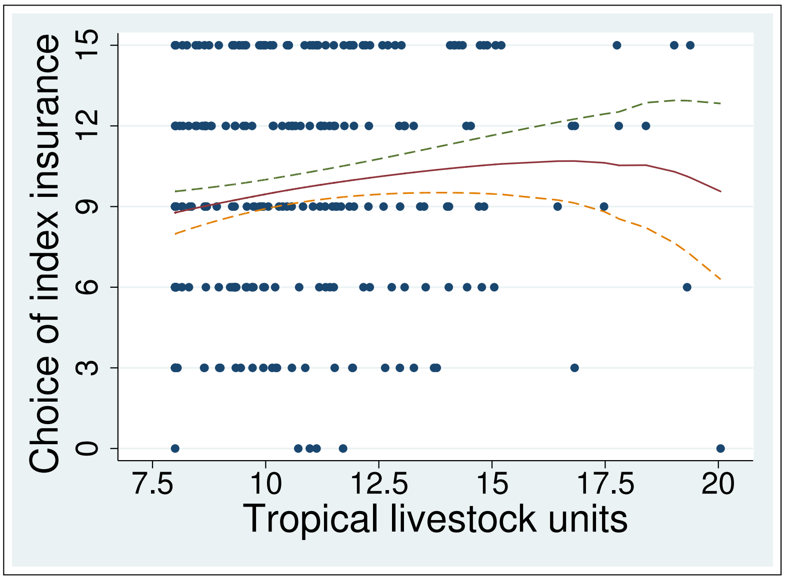
In the following section, we detail, and report the results of, the empirical analysis.

5 Empirical Analysis

The graph in Figure 2 displays the scatter plot of the choices in the individual index insurance decision problem versus tropical livestock units, in addition to

[44] The average household size was greater than five for both the entire ERHS round 7 sample and the sub-sample participating in the experiment.

Figure 2. *Scatter plot and kernel regression: choice in T_{IX} against livestock owned*



Note: The figure shows the point estimate and 95% confidence interval for an Epanechnikov kernel with a bandwidth of 0.8 and trimming of 0.05. The y-axis indicates the amount of insurance chosen (premium choice) in T_{IX}.

a non-parametric kernel regression of choice in this decision problem against tropical livestock units.[45]

The kernel regression line and the corresponding 95% confidence interval in Figure 2 provide some evidence that at low levels of wealth, index insurance take-up is increasing in wealth, and thus decreasing in risk aversion (if, as noted earlier, wealth only impacts take-up in the experiment through risk aversion and preferences satisfy decreasing absolute risk aversion). This is broadly in line with the results of Giné et al. (2008), Cole et al. (2009) and Hill et al. (2011), who find that take-up decreases with risk aversion in their samples of poor households in rural India and Ethiopia.

Figure 2 also provides some indication of non-monotonicity, with index insurance purchase first increasing then decreasing in wealth, as predicted by the theory in Section 2 for expected utility theory and (risk averse) constant relative risk

[45] The kernel regression uses a Nadaraya-Watson estimator, and is a non-parametric technique to estimate the conditional expectation of the choice in the decision problem relative to the livestock units owned. See Li and Racine (2007) for more details on the kernel regression, and Andersen et al. (2008) for use in the analysis of experimental data.

aversion (that is, for expected utility maximizers with preferences satisfying risk aversion and constant relative risk aversion). However, the kernel regression line does not provide conclusive proof of non-monotonicity; it is possible to draw a straight line within the 95% confidence interval (and the confidence interval is quite wide at higher levels of wealth).[46] Index insurance take-up may not be decreasing significantly at high levels of wealth (even though T_{IX} involves actuarially unfair premiums) because, in the sample of experimental subjects, there are no individuals who are so rich as to have very low levels (close to zero) of risk aversion.[47] This is a reasonable assumption, given that the ERHS households are, in general, very poor (Dercon and Krishnan 1998, Porter 2011).

For a more formal analysis of the data, we consider the following reduced form specification:

$$\begin{aligned}
\text{Index Insurance Choice}_i = {} & \beta_0 + \beta_1 \text{Livestock}_i + \beta_2 \text{Livestock}_i^2 \\
& + \beta_3 \text{Std. dev. of consumption}_i + \beta_4 \text{No. of iddir}_i \\
& + \beta_5 \text{Can obtain 100 birr in emergency}_i + \beta_6 \text{If equb}_i \\
& + \beta_7 \text{Household size}_i + \beta_8 \text{Understanding}_i + \beta_9 \text{Quantitative literacy}_i \\
& + \beta_{10} \text{If literate}_i + \beta_{11} \text{Schooling}_i + \beta_{12} \text{Age}_i + \beta_{13} \text{If female}_i \\
& + \beta_{14} \text{If household head}_i + \beta_{15} \text{If farmer}_i \\
& + \beta_{16} \text{Fraction of earnings kept}_i + \epsilon_i \qquad (1)
\end{aligned}$$

where i is the individual subject subscript. For each of OLS and Ordered Probit we consider four specifications: one with only the linear wealth term; one with the linear and squared wealth terms; one with the linear wealth term and all the other explanatory variables; and one with all the terms in specification (1).

Table 2 presents the results for the estimation of these specifications. In these specifications, the subject's choice is the dependent variable – taking values

[46] However, this is only illustrative; a formal test for linearity – such as that developed by Horowitz and Spokoiny (2002) – could be run to confirm this hypothesis in such a context.

[47] As noted in Section 2, only individuals with preferences sufficiently close to risk neutral are expected to have demand for actuarially unfair index insurance close to zero.

1, 2, 3, 4, 5 or 6 – with higher values indicating more insurance purchase (that is, greater premium choice) in T_{IX}. The OLS results are listed in the odd-numbered columns, while the Ordered Probit results are listed in the even-numbered columns.[48] However, since the Ordered Probit results are very similar to OLS results and the dependent variable is an ordinal, ordered response that is not continuous, I focus on analyzing the estimates from the former model.

As would be expected from the kernel regression line (Figure 2), there is no statistically significant linear relationship between demand for index insurance and livestock ownership (columns (2) and (6) of Table 2). However, when the quadratic term is added to the specification, the coefficient estimate of livestock is positive and statistically significant (at the 1% level in the Ordered Probit specifications), and the coefficient estimate of livestock-squared is negative and significant at the 1% level. This, along with the kernel in Figure 2, is suggestive of a hump-shaped non-linear relationship between the take-up of index insurance and wealth.[49]

Additionally, in specifications (4) and (8), the coefficient estimates of the livestock (β_1) and livestock-squared (β_2) terms imply a maximum take-up of index insurance at 13.8 and 13.2 TLUs, respectively.[50] These are marginally greater than the mean TLUs for our experimental sample (10.5), but well below the maximum of 31.5. These results therefore provide evidence for a hump shape, with demand for index insurance first increasing then decreasing with wealth (TLUs), as opposed to just a concave relationship.

[48] Standard errors clustered at the session level and robust to heteroskedasticity are used to construct the t-statistics reported in the tables for all eight specifications in Table 2. This allows for the possibility of correlation between the responses of subjects in the same session, or equivalently that the error term ϵ_i in specification (1) is correlated between individuals in a particular session. The procedures for allowing for clustering also allow heteroskedasticity between and within clusters, as well as autocorrelation within clusters (Andersen et al. 2008); however, these procedures still assume that the error is uncorrelated between subjects across different sessions. The results of the hypothesis tests do not differ markedly when non-robust and non-clustered standard errors are utilized, indicating that heteroskedasticity and correlation of the error term between individuals in a particular session may not be major concerns.

[49] These results are also indicative of a hump-shaped non-linear relationship between take-up and risk aversion, assuming that the effect of wealth on take-up in the lab experiment is only through risk aversion and preferences satisfy decreasing absolute risk aversion.

[50] For $\beta_1 > 0$ and $\beta_2 < 0$, the quadratic reaches a maximum when livestock equals $-\frac{\beta_1}{2\beta_2}$.

Table 2. *Correlates of index insurance take-up in* T_{IX}

Variables	(1) OLS	(2) O. Probit	(3) OLS	(4) O. Probit	(5) OLS	(6) O. Probit	(7) OLS	(8) O. Probit
Tropical livestock units	0.0346	0.0222	0.517***	0.452***	-0.0307	-0.0294	0.401**	0.412***
	(0.600)	(0.513)	(4.830)	(3.972)	(0.510)	(0.594)	(2.548)	(2.653)
Livestock squared			-0.0180***	-0.0164***			-0.0148***	-0.0156***
			(5.570)	(3.551)			(3.235)	(2.806)
Std. dev. of consumption					0.000224	0.000173	0.000137	0.000112
					(0.780)	(0.788)	(0.479)	(0.500)
No. of iddir					0.0724	0.0431	0.0338	0.00716
					(0.845)	(0.707)	(0.395)	(0.117)
Can obtain 100 Birr					0.749***	0.607***	0.700***	0.574***
					(3.073)	(3.337)	(2.949)	(3.228)
If equb					-0.14	-0.086	-0.166	-0.108
					(0.528)	(0.423)	(0.629)	(0.527)
Understanding					-0.0256	-0.103	0.0644	-0.0432
					(0.031)	(0.155)	(0.078)	(0.064)
Quantitative literacy					0.579	0.539	0.594	0.55
					(0.912)	(1.150)	(0.970)	(1.203)
If literate					0.511*	0.425**	0.487*	0.419*
					(1.821)	(1.969)	(1.731)	(1.927)
Schooling obtained					-0.000859	-0.000994	0.00775	0.00591
					(0.032)	(0.048)	(0.299)	(0.294)
Age					0.00322	0.00216	0.00217	0.000885
					(0.434)	(0.383)	(0.301)	(0.161)
If female					0.525*	0.377*	0.493	0.351
					(1.851)	(1.651)	(1.700)	(1.499)
If household head					-0.699***	-0.591***	-0.631***	-0.530***
					(3.153)	(3.400)	(2.822)	(2.990)
If farmer					0.413	0.311	0.397	0.301
					(1.562)	(1.467)	(1.531)	(1.438)
Household size					0.00937	0.00545	-0.0109	-0.0137
					(0.222)	(0.164)	(0.267)	(0.421)
Fraction of earnings kept					-0.429*	-0.391**	-0.409*	-0.380**
					(1.886)	(2.145)	(1.829)	(2.104)
Constant	3.819***		0.844		2.738**		-0.0292	
	(6.255)		(1.031)		(2.547)		(0.025)	
Threshold 1		-1.673***		0.862		-1.233		1.474
		(3.463)		(1.253)		(1.398)		(1.553)
Threshold 2		-0.786*		1.810**		-0.284		2.471**
		(1.681)		(2.554)		(0.327)		(2.572)
Threshold 3		-0.202		2.408***		0.38		3.141***
		(0.452)		(3.260)		(0.458)		(3.205)
Threshold 4		0.357		2.978***		0.992		3.759***
		(0.758)		(4.022)		(1.154)		(3.684)
Threshold 5		0.910**		3.542***		1.618*		4.393***
		(2.013)		(4.673)		(1.935)		(4.317)
Order Controls	No	No	No	No	Yes	Yes	Yes	Yes
Enumerator Controls	No	No	No	No	Yes	Yes	Yes	Yes
Location Controls	No	No	No	No	Yes	Yes	Yes	Yes
Observations	254	254	254	254	246	246	246	246
R-squared	0.004		0.049		0.188		0.210	
Log-likelihood		-423.1		-416.3		-383.7		-379.4

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$
Absolute values of t-statistics in parentheses.
Dependent variable is the respondent's choice in the index insurance decision problem (T_{IX}).

The results are consistent with expected utility theory and (risk averse) constant relative risk aversion, in accordance with the theory detailed in Section 2; they contradict the predictions of Giné et al.'s (2008) mean variance model and those of technology adoption models. Thus, the results indicate that the theoretical model presented in Section 2, rather than these alternative models, may be more suitable for describing the shape of index insurance take-up. In conformity with the theory outlined in Section 2, index insurance take-up is "rationally" low (that is, due to a fundamental lack of desirability of the product) for the poorest (most risk averse) and wealthiest (least risk averse) individuals, as a result of basis risk[51] and actuarially unfair premiums respectively.[52,53]

While Giné et al. (2008), Cole et al. (2009) and Hill et al. (2011) also find that the most risk averse individuals in their samples have the lowest take-up of index insurance, they find that take-up decreases monotonically with risk aversion and increases monotonically with wealth.[54] Thus, in effect, all three studies find evidence against Giné et al.'s (2008) mean variance model. They attribute the low observed take-up, particularly by the poorest and most risk averse individuals, to barriers associated with the adoption of a new technology such as poor understanding, lack of trust or doubts about the credibility of plans (of the insurance vendor), credit constraints, unwillingness to experiment or irrationality on the part of rural consumers; they conclude that take-up of index

[51] Hill et al. (2011) find that basis risk significantly reduces the demand for index insurance.

[52] It is important to acknowledge the possibility that some of the explanatory variables in the specifications, such as wealth and background risk, may be endogenous (see Chapter 2 for more details). However, since the primary aim of this paper is to shed light on the relationship between index insurance take-up and wealth, and not to study asset integration, the issue of endogeneity is less salient for inference in this case. Further, Giné et al. (2008), Cole et al. (2009) and Hill et al. (2011) do not account for this possible endogeneity in their analyses of the correlates of index insurance take-up. Along the lines of these studies, this chapter focuses on evaluating the correlates of index insurance take-up (and in particular, investigating the relationship between take-up and wealth), rather than establishing the causal impact of various factors on risk preferences (which is the focus of Chapter 2).

[53] We cannot definitively rule out other explanations for the observed hump-shaped relationship between index insurance take-up and wealth. However, the results indicate that we cannot reject the theoretical model, and the factors driving it. The Ethiopian experimental data is consistent with the theoretical model outlined in this chapter, but may be consistent with other models as well. The results of this analysis supply important implications for providers of index insurance, and further research on this issue is required.

[54] Similar results are obtained by Lybbert et al. (2010) using experimental data from Peru and Kenya.

insurance is best described by technology adoption models.[55] However, our results indicate that the low take-up observed, particularly among the poorest (and most risk averse) individuals, in these studies may be a result of rational choice rather than any of these barriers.[56]

It is also important to note that while we find the non-linear relationship between index insurance take-up and wealth to be consistent with the hump-shape theoretical prediction in Section 2, the levels of take-up observed in T_{IX} are significantly greater than those predicted by Theorem 4 of Clarke (2011), for expected utility maximizers with preferences satisfying risk aversion and decreasing absolute risk aversion.[57] In this chapter we analyze only the shape, not the level, of insurance purchase in the index insurance decision problem; additionally, our theoretical model is based on expected utility theory and constant relative risk aversion preferences. This is in line with the work of Giné et al. (2008) and Hill et al. (2011), who do not make theoretical predictions about the level of index insurance take-up but rather focus on predicting, and empirically testing for, the shape of the relationship of take-up with respect to factors such as wealth and risk aversion, within the framework of expected utility theory. Section 6 in Chapter 1 contains an analysis of the level of index insurance take-up in the Ethiopian experiment.[58] Further, as noted earlier, analyzing the correlates of index insurance take-up is crucial, as it can help identify, and address, the causes of the low take-up observed in pilot programs (Giné et al. 2008).

[55] Additionally, Hill and Nobles (2011) also find very low take-up of real weather indexed insurance securities offered to farmers in rural Ethiopia, and attribute this to liquidity constraints.

[56] Additionally, as acknowledged by Hill et al. (2011), even though index insurance is a new and innovative financial product, decisions regarding its purchase may not be perfectly governed by technology adoption models – this is because unlike agricultural technologies, insurance is not a physical product and it only provides benefits in certain years (which may be some time after its purchase).

[57] Relatively high take-up rates of index insurance have also been observed in developing country lab experiments by Hill et al. (2009) (weather indexed insurance) and Lybbert et al. (2010) (area yield indexed insurance), even though pilot programs in the field have found relatively low levels of take-up.

[58] Though the shape of the relationship between index insurance take-up and wealth is found to be in line with expected utility theory and constant relative risk aversion preferences, the level of take-up in the Ethiopian experiment is not, as noted in Section 6 of Chapter 1. Further research is required to shed light on this apparent divergence.

There is also some evidence that literate subjects purchase more index insurance – the coefficient estimate of the 'If literate' variable is positive and statistically significant at the 5% or 10% level in all specifications. Unlike Cole et al. (2009), however, we do not find that other proxies for cognitive ability, such as quantitative literacy and understanding, significantly increase the take-up of index insurance. Further, in line with work of Giné et al. (2008), we find no significant relationship between formal education and take-up.

Therefore, there is only limited evidence for the hypothesis of Giné et al. (2008) and Cole et al. (2009) that limited cognition and poor understanding of index insurance are important causes of the low take-up in pilot programs. In line with this observation, Hill and Nobles (2011) find that weather indexed insurance products offered to farmers in Southern Ethiopia in both an experimental setting and in real life were well understood, and in particular the participants involved were able to comprehend the basis risk associated with the policies – as a result, they note that such products may be appropriate for populations with relatively low quantitative and financial literacy. Similarly, Hill et al. (2011) also find evidence that most participants in the ERHS understood a hypothetical index insurance product and the associated basis risk. Thus, improving understanding of the product (a strategy suggested by Lybbert et al. 2010) may not lead to significant increases in the take-up of index insurance, given that lack of understanding does not seem to be a major obstacle in its take-up. These results may also indicate that what really matters for the take-up of index insurance is basic literacy, rather than the years of formal schooling obtained or quantitative literacy.

We also find some evidence, in line with the work of Cole et al. (2009), that a household's experience with other forms of insurance (such as informal insurance) and the size of its informal risk-sharing network are associated with greater take-up of index insurance. The coefficient estimate of the dummy indicating whether the subject's household can obtain 100 birr in the case of an emergency is positive and significant at the 1% level in all specifications.

However, there is no statistically significant relationship between take-up and iddir or equb membership, indicating that membership in these networks may not be associated with greater index insurance purchase. On the other hand, Giné et al. (2008) find that households in rural India which are members of the local village council are significantly more likely to purchase index insurance, and conclude that membership in social networks is a key determinant of index insurance take-up. The results obtained by Hill et al. (2011) on the relationship between access to alternative insurance arrangements and index insurance take-up are mixed – while they find that the size of a subject's informal risk-sharing network has a positive association with index insurance take-up, they also find that the estimated effect of the same 100 birr dummy variable on the take-up of the hypothetical insurance contract is negative and statistically significant. Thus, further research is required to determine what, if any, are the channels through which access to alternative insurance mechanisms and membership in social groups affect the take-up of index insurance.

Further, the coefficient estimates of the inter-temporal standard deviation of consumption are statistically insignificant and extremely small in magnitude (Columns (6) and (8)), lending credence to the hypothesis that background risk does not significantly impact subjects' choices, and thus the take-up of index insurance, in decision problem T_{IX}.[59] This is not in line with Gollier and Pratt's (1996) risk vulnerability hypothesis, according to which higher background risk should be associated with greater (indirect) risk aversion, and thus should affect insurance take-up in the experiment as well.[60] Lusk and Coble (2008), analyzing the influence of experimentally-induced background risk among university subjects in the United States, reject the risk vulnerability hypothesis of Gollier and Pratt (1996) and find only a small impact of background risk on risk

[59] Note that Giné et al. (2008) and Cole et al. (2009) do not control for background risk in their analyses of the correlates of index insurance take-up in rural India.

[60] Strictly speaking, Gollier and Pratt's (1996) risk vulnerability hypothesis, applied to the context of index insurance take-up, implies that the relationship between take-up and background risk should be hump-shaped as well. To explore this, I conducted all the specifications in this section with the inclusion of a squared background risk term (in addition to the level term) – however, there are no substantive differences in the results.

preferences. Additionally, Alessie et al. (2002), analyzing the portfolio structure of Dutch households using survey data, find no significant relationship between the demand for risky assets and background income uncertainty. These results are also in line with those obtained in Chapter 2 using data from the same Ethiopian experiment, where I found no significant relationship between background risk and risk attitudes. However, these results are in contrast to those of Harrison et al. (2007) and Guiso and Paiella (2008), who find that experimentally-induced and non-experimental background risk, respectively, significantly increase risk aversion.[61]

The results in Table 2 also show that subjects who were not household heads tended to purchase more index insurance. This may be because those subjects who were answerable to their household heads – and probably had to give most (or all) of their experimental earnings to their household heads – were averse to returning with little (or no) money from the experiment (other than the participation fee) and thus were more likely to choose greater insurance purchase. Also in line with this explanation, the results show that the fraction of experimental earnings kept by the subject has a negative and statistically significant relationship with index insurance take-up. Additionally, unlike Hill et al. (2011), we do not find strong evidence that age and gender are associated with index insurance take-up. Further, it should also be noted that none of the statistically significant coefficient estimates in Table 2 differ in sign between the OLS and Ordered Probit models. This gives us some confidence in the robustness of the OLS results obtained in this section as well.

6 Robustness Checks

The results of the previous section withstand common robustness checks. One concern might be that outliers in the ERHS data on livestock ownership are

[61] The results of Guiso and Paiella (2008), however, are based on a single hypothetical survey question – as noted earlier, the use of non-incentivized, hypothetical tasks could yield misleading results.

driving the results involving this variable; this is often a worry when using survey data (Chambers and Skinner 2003). However, all the specifications in Section 5 produce similar results when the logarithm of tropical livestock units is used instead of the level. This indicates that the results obtained in the previous section are probably not driven by outliers in the wealth measure.

If the analysis is restricted to include only those subjects who played T_{IX} as the first of their three decision problems in the session, or only those subjects who answered all the confirmation of understanding questions correctly[62], the results remain substantively the same. This suggests that neither order effects nor subjects with low levels of understanding are driving the results (Harbaugh et al. 2002).

Tropical livestock units are used as the primary measure of wealth in our analysis, as is common in studies using ERHS data. However, this measure may not appropriately discriminate between wealthier and poorer participants, since the households of different subjects may hold wealth in different forms. We therefore construct a composite wealth measure, using the ten main asset questions in the ERHS, and repeat the analysis conducted in Section 5 for the index insurance decision problem T_{IX} using this wealth measure. This wealth measure (index) is constructed to be the first principal component arising from a principal components analysis (PCA) of these ten asset variables for the entire ERHS sample (see Table 3).[63] Filmer and Pritchett (2001), using survey data from rural India to construct and analyze a similar wealth index, find that a PCA wealth index generated from the first principal component using discreet asset ownership indicators provides a reliable proxy for the long-run economic status of households. Additionally, Vyas and Kumaranayake (2006) note that wealth measures based on asset data – rather than income or

[62] These questions are reported in Table 1.

[63] Moser and Felton (2007) and Howe et al. (2008) provide a detailed analysis of PCA as a tool for constructing wealth indices, including some of the advantages and disadvantages. PCA is not scale invariant and so the analysis in this section is conducted using the correlation matrix. This is equivalent to using standardized variables.

Table 3. *Weights assigned to each indicator in the PCA wealth index*

Item	Item weight in PCA wealth index
Tropical livestock units	0.362
Land	0.295
If own bank account	0.268
If own hoe	0.272
If own plough	0.325
If own radio	0.397
If own gold	0.339
If own bicycle	0.214
If own cell phone	0.338
If currently storing any crop	0.313

consumption data – involve lower measurement error. Such PCA indices have been used extensively in published studies as measures of wealth in the context of developing countries (for example, Filmer and Pritchett 2001, Schellenberg et al. 2003), including in analyses involving data from Ethiopia (for example, Vyas and Kumaranayake 2006). Table 3 presents the weights from the principal eigenvector, which are used to construct the index.[64,65] As might be expected, all weights have the same sign, indicating that they are jointly positively correlated. Figure 3(a), which utilizes data from the experimental sample, shows that the resulting PCA wealth index is highly correlated with tropical livestock units, suggesting that tropical livestock units are a good proxy for wealth in our sample sites.

Figure 3(b) and Table 4 present the scatter plots, non-parametric kernel regressions and correlates of index insurance take-up when this PCA wealth measure is used in place of tropical livestock units, and may be directly compared with Figure 2 and Table 2 respectively. As can be seen from the figure and table, all results remain substantively the same when TLUs are replaced with the PCA wealth index. In particular, in the Ordered Probit quadratic specifications reported in Columns (4) and (8) of Table 4, the coefficient estimate of the PCA wealth term is positive and statistically significant, while that of the squared term is

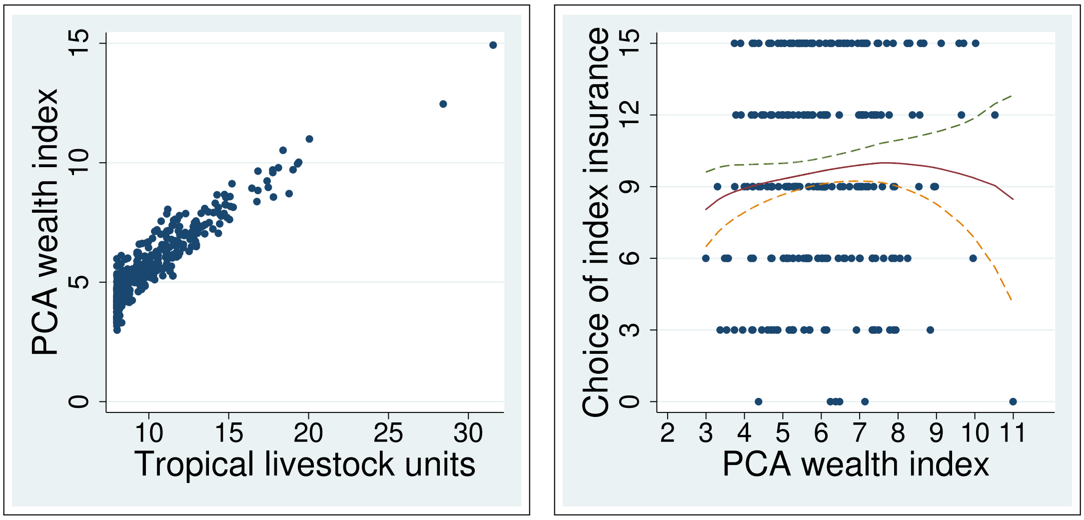
[64] Note that of the ten variables used to construct the PCA wealth index, all except tropical livestock units and land owned (measured in hectares) are binary dummy variables which equal one if the subject owned that particular asset and zero otherwise.

[65] See Section 3.2 in Chapter 3 for more details on the construction of a PCA wealth index.

negative and statistically significant. Further, these estimates suggest that index insurance purchase is highest when the PCA wealth index is 7.0, again within one standard deviation of the mean PCA wealth index for our experimental sample (approximately 6) and well below the maximum (30.9). This provides strong evidence that demand for index insurance is hump-shaped – not just increasing concave – in wealth.

Figure 3. *Results from the PCA wealth index*

(a) PCA wealth versus tropical livestock units (b) Scatter plot and kernel regression for T_{IX}



Note: Figure (b) presents the point estimate and 95% confidence interval for an Epanechnikov kernel with a bandwidth of 0.8 and trimming of 0.05.

Finally, although not reported in this chapter, there was another index insurance decision problem played by some subjects, similar to T_{IX} but where individuals were randomly placed in risk-pooling pairs – that is, the group index insurance problem T_{GX} described in Chapter 1. The results of Section 5 hold if we add the decisions made in this paired index insurance problem.

7 Conclusion

Lab and field experiments are complementary (Falk and Heckman 2009, Harrison and List 2004). Lab experiments conducted in the field with non-standard subject pools of developing country inhabitants provide an important middle ground

Table 4. *Correlates of index insurance take-up in T_{IX} with PCA wealth measure*

Variables	(1) OLS	(2) O. Probit	(3) OLS	(4) O. Probit	(5) OLS	(6) O. Probit	(7) OLS	(8) O. Probit
PCA Wealth	0.0466	0.0299	0.918***	0.710**	-0.0738	-0.0681	0.793*	0.721*
	(0.484)	(0.420)	(2.911)	(2.499)	(0.587)	(0.662)	(1.793)	(1.776)
PCA Wealth squared			-0.0652**	-0.0512**			-0.0593*	-0.0544*
			(2.546)	(2.166)			(1.927)	(1.844)
Std. dev. of consumption					0.000169	0.000123	0.000109	0.0000808
					(0.605)	(0.581)	(0.400)	(0.390)
No. of iddir					0.118	0.0774	0.0679	0.0332
					(1.352)	(1.200)	(0.791)	(0.522)
Can obtain 100 Birr					0.745***	0.604***	0.664***	0.540***
					(3.297)	(3.448)	(2.797)	(2.935)
If equb					-0.0877	-0.0392	-0.126	-0.0738
					(0.323)	(0.190)	(0.447)	(0.341)
Understanding					-0.365	-0.36	-0.368	-0.372
					(0.424)	(0.521)	(0.420)	(0.526)
Quantitative literacy					0.553	0.535	0.476	0.468
					(0.844)	(1.098)	(0.731)	(0.954)
If literate					0.605**	0.493**	0.645**	0.535**
					(2.179)	(2.249)	(2.275)	(2.359)
Schooling obtained					0.00143	0.00116	0.000277	-0.0000969
					(0.054)	(0.057)	(0.011)	(0.005)
Age					0.00272	0.00182	0.0012	0.000288
					(0.396)	(0.350)	(0.174)	(0.055)
If female					0.498	0.359	0.477	0.344
					(1.699)	(1.526)	(1.618)	(1.441)
If household head					-0.723***	-0.614***	-0.660**	-0.557***
					(3.036)	(3.274)	(2.757)	(2.909)
If farmer					0.369	0.282	0.367	0.284
					(1.309)	(1.245)	(1.298)	(1.234)
Household size					0.00732	0.00499	-0.0106	-0.0114
					(0.168)	(0.142)	(0.250)	(0.338)
Fraction of earnings kept					-0.431*	-0.389**	-0.421*	-0.379**
					(1.934)	(2.197)	(1.883)	(2.130)
Constant	3.906***		1.168		3.095**		0.392	
	(6.790)		(1.140)		(2.786)		(0.276)	
Threshold 1		-1.717***		0.357		-1.498*		0.912
		(3.832)		(0.444)		(1.656)		(0.759)
Threshold 2		-0.834*		1.282		-0.55		1.896
		(1.878)		(1.488)		(0.616)		(1.520)
Threshold 3		-0.267		1.859**		0.111		2.565**
		(0.637)		(2.096)		(0.129)		(2.018)
Threshold 4		0.304		2.435***		0.737		3.194**
		(0.702)		(2.788)		(0.841)		(2.486)
Threshold 5		0.856**		2.991***		1.367		3.826***
		(2.046)		(3.292)		(1.596)		(2.942)
Order Controls	No	No	No	No	Yes	Yes	Yes	Yes
Enumerator Controls	No	No	No	No	Yes	Yes	Yes	Yes
Location Controls	No	No	No	No	Yes	Yes	Yes	Yes
Observations	244	244	244	244	237	237	237	237
R-squared	0.002		0.028		0.198		0.213	
Log-likelihood		-407.2		-403.6		-368.8		-366.2

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$
Absolute values of t-statistics in parentheses.
Dependent variable is the respondent's choice in the index insurance decision problem (T_{IX}).

between the lab and the field, offering a significant degree of control while increasing external validity (Harrison et al. 2007). The Ethiopian experiment – and the index insurance decision problem within it – provides fertile ground to analyze the demand for index insurance products, which in recent years have emerged as a popular medium for providing insurance to poor farmers in developing countries. The experiment facilitates the rigorous analysis of the demand for index insurance – a product for which the contractual index is not perfectly correlated with the loss – as the losses and index are generated in a laboratory setting by a known randomization device with objective joint probability distribution. By contrast, it is impossible to obtain a precise objective estimate of the joint probability distribution of losses and index claim payments using field data, which makes it difficult to draw conclusions and make clear normative statements about observed demand for index insurance; additionally, the laboratory setting enables isolation from the credit and liquidity constraint effects of wealth on index insurance take-up.

In this chapter, we present a theory which outlines the shape of the relationship between optimal index insurance take-up and wealth, and then empirically test it using the experimental data from rural Ethiopia. In our lab experiment, we find that the demand for index insurance from rural Ethiopians is hump-shaped in wealth. That is, the demand is first increasing then decreasing in wealth, with the poorest participants (who are expected to be the most risk averse) and the richest participants (who are expected to be the least averse to risk) in the experiment having the lowest demand, while the highest demand is from subjects with intermediate levels of wealth. This shape of the demand for index insurance is consistent with expected utility theory and (risk averse) constant relative risk aversion – in line with the theory presented in Section 2 – which implies that the take-up of index insurance is "rationally" low for the poorest (most risk averse) and wealthiest (least risk averse) individuals, due to basis risk and actuarially unfair premiums respectively. Further, this shape contradicts the predictions of the "neoclassical" mean variance model proposed by Giné et al. (2008) to describe the demand for index insurance, as well the predictions of technology adoption models which Hill et al. (2011) believe to be appropriate for informing the take-up of index insurance. Thus, the results indicate that the theoretical model presented in Section 2, rather than these models, may be more suitable for describing the shape of index insurance take-up.

Giné et al. (2008) Cole et al. (2009) and Hill et al. (2011), on the other hand, find evidence that index insurance take-up increases monotonically with wealth and decreases monotonically with risk aversion. They attribute this finding, including the low take-up observed among the poorest and most risk averse farmers, to barriers associated with the adoption of a new technology, such as lack of understanding, credit constraints, unwillingness to experiment with new products or poor decision-making on the part of rural consumers. However, the results in this chapter indicate that this may instead be a result of rational (or optimal) choice due to a fundamental lack of desirability of the product – caused by basis risk and actuarially unfair premiums – rather than any of these factors.

If demand for index insurance is "rationally" low for the most risk averse individuals – who are also likely to be the poorest – it implies that the measures outlined by Giné et al. (2008) and Cole et al. (2009) to address credit constraints and "behavioural" impediments to index insurance take-up may not be effective in increasing take-up by the poorest.[66] Further, our results contradict the conclusion of Giné et al. (2008) and Cole et al. (2009) that the low take-up rates among the poorest do not reflect a lack of demand (or fundamental lack of desirability) of the product and may on the other hand reflect the "normal pattern of diffusion of a new product" (in accordance with the literature on technology adoption), which would be expected to accelerate over time.

To increase index insurance purchase by the poorest individuals, other provisions (such as government-provided safety nets) may be required – in combination with

[66] These proposed measures include, for example, the minimization of transaction costs associated with index insurance, rapid payouts to mitigate liquidity constraints and the combination of insurance contracts with short-term loans to address credit constraints.

index insurance contracts – to mitigate or protect against the negative basis risk associated with indexed products. Hill et al. (2011) note that the design of existing index insurance products is far from perfect, and research on this design – with a focus on reducing basis risk – as well as on methods to increase its take-up is essential.

One method for increasing the take-up of index insurance, proposed by Dercon et al. (2011), is to market and sell the products to pre-existing risk-sharing groups (such as iddirs in Ethiopia). Using a theoretical model and evidence from a randomized experiment in Ethiopia, they find that take-up of this product is substantially greater when it is marketed to insurance groups rather than individuals. Further, they find that this is primarily because group members can share the basis risk associated with index insurance contracts, since all basis risk is not perfectly correlated among group members; thus, while informal insurance arrangements within the group help smooth consumption in the face of idiosyncratic risk (hence mitigating the effect of basis risk on each of the members), the index insurance contracts provide a useful method to insure the group as a whole, particularly against covariate weather shocks. Similar evidence is found by Hill et al. (2011). Therefore, selling index insurance to such groups could significantly increase its take-up (including by the most risk averse), while reducing transaction costs and providing insurance against covariate shocks (Dercon et al. 2011).

Hill and Nobles (2011) argue that there is substantial heterogeneity in rainfall risk even for farmers within a small geographical area – thus, they find that offering a range of flexible weather indexed insurance products rather than traditional individual pre-packaged index insurance contracts facilitates improved management of basis risk and increases the take-up of index insurance. Further research into these alternatives is essential. However, Hill and Nobles (2011) also note that basis risk will continue to be an issue for any weather indexed insurance product, and further research is also required to determine how index insurance can be integrated with the other risk-coping mechanisms currently used by rural inhabitants, such as group-based risk sharing as well as informal credit and gift-giving. Alternatively, insurance providers need to design new insurance products – which involve little (or no) basis risk – that are attractive to very risk averse individuals.

The results in this chapter indicate that there is no statistically significant association between background risk and index insurance take-up. This result is not in line with Gollier and Pratt's (1996) risk vulnerability hypothesis, but is consistent with the results in Chapter 2 as well as those obtained by Alessie et al. (2002) and Lusk and Coble (2008). We also find that basic literacy, rather than years of formal schooling or quantitative literacy, is associated with greater take-up of index insurance. Further, we find some evidence that a household's experience with other forms of insurance (such as informal insurance) and the size of its informal risk-sharing network increase the take-up of index insurance.

Development economists and insurance practitioners are rightly optimistic about the potential for index insurance to substantially increase welfare for many of the world's rural poor (Skees et al. 1999, Karlan and Morduch 2009). In particular, weather indexed insurance provides a promising avenue for rural households in developing countries to lower their risk exposure, since weather variability is often cited as the most important source of risk in these regions (Cole et al. 2009). However, voluntary purchase of unsubsidized products remains low. All too often this low demand is labeled a "puzzle" and attributed to poor "behavioural" decision-making on the part of farmers, without proper consideration of whether the product provides good value to rational farmers. As noted by Giné et al. (2008), microinsurnace markets are expanding swiftly and innovative products such as index insurance are growing in popularity with policymakers; however, academic research on microinsurance – and index insurance in particular – is still in its infancy, and many crucial questions remain unanswered. Further research on this topic is therefore critical.

Regardless of how precisely insurance for the poor develops, it seems important for social scientists to be engaged with normative questions of interest to insurance providers, consumers and regulators. On their own, field experiments seem unlikely to advance our understanding of key questions such as how to design good products or whether clients make good financial decisions (that is, optimal or rational decisions). To contribute to these questions we will also need normative theory and carefully designed lab experiments.

References

Alessie, RJM, S. Hochgürtel, and A.H.O. Soest, "Household portfolios in the Netherlands," in L. Guiso, M. Haliassos, and T. Jappelli, eds., *Household Portfolios*, MIT Press, 2002, pp. 341–388.

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

Bandiera, O. and I. Rasul, "Social Networks and Technology Adoption in Northern Mozambique*," *The Economic Journal*, 2006, *116* (514), 869–902.

Barrett, C. and J. McPeak, "Poverty traps and safety nets," *Poverty, Inequality and Development*, 2006, pp. 131–154.

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

Chambers, R.L. and CJ Skinner, *Analysis of survey data*, John Wiley & Sons Inc, 2003.

Charness, G. and A. Viceisza, "Comprehension and risk elicitation in the field: Evidence from rural Senegal," *IFPRI Discussion Paper*, 2012.

Charness, Gary, Uri Gneezy, and Alex Imas, "Experiential methods: Eliciting risk preferences," *Journal of Economic Behavior & Organization*, 2013.

Clarke, D.J., "A Theory of Rational Demand for Index Insurance," Department of Economics Discussion Paper Series 572, University of Oxford 2011.

Cole, Shawn A., Xavier Giné, Jeremy B. Tobacman, Petia B. Topalova, Robert M. Townsend, and James I. Vickery, "Barriers to Household Risk Management: Evidence from India," Working Paper 09-116, Harvard Business School 2009.

Dave, C., C.C. Eckel, C.A. Johnson, and C. Rojas, "Eliciting risk preferences: When is simple better?," *Journal of Risk and Uncertainty*, 2010, pp. 1–25.

Dercon, S., "Growth and shocks: evidence from rural Ethiopia," *Journal of Development Economics*, 2004, *74* (2), 309–329.

—— **and P. Krishnan**, "Changes in poverty in rural Ethiopia 1989-1995: measurement, robustness tests and decomposition," *CSAE Working Paper Series*, 1998, *98.7*.

Dercon, Stefan, Ruth Vargas Hill, Ingo Outes-Leon, A Bayrou, D Clarke, and A Seyoum-Taffesse, "Offering rainfall insurance to informal insurance groups: evidence from a field experiment in Ethiopia," Technical Report, Mimeo 2011.

Diamond, D.W., "Financial intermediation and delegated monitoring," *The Review of Economic Studies*, 1984, *51* (3), 393–414.

Doherty, N.A. and H. Schlesinger, "Rational insurance purchasing: consideration of contract nonperformance," *The Quarterly Journal of Economics*, 1990, *105* (1), 243–253.

Eeckhoudt, Louis, Christian Gollier, and Harris Schlesinger, "The risk-averse (and prudent) newsboy," *Management Science*, 1995, *41* (5), 786–794.

Fafchamps, M., *Rural Poverty, Risk and Development*, Edward Elgar Publishing, 2003.

Falk, A. and J.J. Heckman, "Lab experiments are a major source of knowledge in the social sciences," *Science*, 2009, *326* (5952), 535.

Feder, G., R.E. Just, and D. Zilberman, "Adoption of Agricultural Innovations in Developing Countries: A Survey," *Economic Development and Cultural Change*, 1985, *33* (2), 255.

Filmer, D. and L.H. Pritchett, "Estimating Wealth Effects Without Expenditure Data – Or Tears: An Application To Educational Enrollments In States Of India," *Demography*, 2001, *38* (1), 115–132.

Giné, X. and D. Yang, "Insurance, credit, and technology adoption: Field experimental evidencefrom Malawi," *Journal of Development Economics*, 2009, *89* (1), 1–11.

——, R. Townsend, and J. Vickery, "Statistical Analysis of Rainfall Insurance Payouts in Southern India," *American Journal of Agricultural Economics*, 2007, *89* (5), 1248–1254.

——, ——, and ——, "Patterns of rainfall insurance participation in rural India," *The World Bank Economic Review*, 2008, *22* (3), 539.

Gollier, C. and J.W. Pratt, "Risk vulnerability and the tempering effect of background risk," *Econometrica: Journal of the Econometric Society*, 1996, pp. 1109–1123.

Guiso, L. and M. Paiella, "Risk aversion, wealth, and background risk," *Journal of the European Economic Association*, 2008, *6* (6), 1109–1150.

Harbaugh, W.T., K. Krause, and L. Vesterlund, "Risk attitudes of children and adults: Choices over small and large probability gains and losses," *Experimental Economics*, 2002, *5* (1), 53–84.

Harrison, G.W. and J.A. List, "Field experiments," *Journal of Economic Literature*, 2004, *42* (4), 1009–1055.

——, ——, and C. Towe, "Naturally occurring preferences and exogenous laboratory experiments: A case study of risk aversion," *Econometrica*, 2007, *75* (2), 433–458.

Hazell, P., "The appropriate role of agricultural insurance in developing countries," *Journal of International Development*, 1992, *4* (6), 567–581.

Hill, R. V. and M. Nobles, "Flexible insurance for heterogenous farmers: results from a small scale pilot in Ethiopia," *IFPRI discussion paper 01092*, 2011.

Hill, RV, A Viceisza, and J Deustua-Rossel, "The welfare and behavioral impact of insurance provision in rural Ethiopia," *Report submitted to USAID*, 2009.

Hill, R.V., J. Hoddinott, and N. Kumar, "Adoption of weather index insurance: Learning from willingness to pay among a panel of households in rural Ethiopia," *IFPRI discussion paper 01088*, 2011.

Hoddinott, J., S. Dercon, and P. Krishnan, "Networks and Informal Mutual Support in 15 Ethiopian Villages," *Institutional Economics Perspectives on African Agricultural Development*, 2005, p. 273.

Holt, C.A. and S.K. Laury, "Risk aversion and incentive effects," *American Economic Review*, 2002, *92* (5), 1644–1655.

Horowitz, Joel L and Vladimir G Spokoiny, "An adaptive, rate-optimal test of linearity for median regression models," *Journal of the American Statistical Association*, 2002, *97* (459), 822–835.

Howe, L.D., J.R. Hargreaves, and S.R.A Huttly, "Issues in the Construction of Wealth Indices for the Measurement of Socio-Economic Position in Low-Income Countries," *Emerging Themes in Epidemiology*, 2008, *5* (3).

Jalan, J. and M. Ravallion, "Transient poverty in postreform rural China," *Journal of Comparative Economics*, 1998, *26* (2), 338–357.

Kamanou, G. and J. Morduch, "8. Measuring Vulnerability to Poverty," *Insurance Against Poverty*, 2004, *1* (9), 155–176.

Karlan, D. and J. Morduch, "Access to Finance: Credit Markets, Insurance, and Saving," in Dani Rodrik and Mark Rosenzweig, eds., *Handbook of Development Economics*, Vol. 5, Amsterdam: North Holland, 2009, chapter 71.

Li, Q. and J.S. Racine, *Nonparametric econometrics: Theory and practice*, Princeton University Press Princeton, NJ, 2007.

Lusk, J.L. and K.H. Coble, "Risk aversion in the presence of background risk: Evidence from an economic experiment," *Research in Experimental Economics*, 2008, *12*, 315–340.

Lybbert, Travis J, Francisco B Galarza, John McPeak, Christopher B Barrett, Stephen R Boucher, Michael R Carter, Sommarat Chantarat, Aziz Fadlaoui, and Andrew Mude, "Dynamic Field Experiments in Development Economics: Risk Valuation in Morocco, Kenya, and Peru," *Agricultural and Resource Economics Review*, 2010, *39*, 1–17.

Moser, C. and A. Felton, "The construction of an asset index measuring asset accumulation in Ecuador," *Chronic Poverty Research Centre Working Paper*, 2007.

Porter, C., "Examining the impact of idiosyncratic and covariate shocks on Ethiopian households' consumption and income sources," Technical Report 2011.

Pratt, J.W., "Risk Aversion in the Small and in the Large," *Econometrica: Journal of the Econometric Society*, 1964, *32* (1), 122–136.

Quiggin, J., "Background risk in generalized expected utility theory," *Economic Theory*, 2003, *22* (3), 607–611.

Rosenzweig, M.R. and H.P. Binswanger, "Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments," *The Economic Journal*, 1993, *103* (416), 56–78.

Schellenberg, J.A., C.G. Victora, A. Mushi, D. de Savigny, D. Schellenberg, H. Mshinda, and J. Bryce, "Inequities among the very poor: health care for children in rural southern Tanzania," *The Lancet*, 2003, *361* (9357), 561–566.

Skees, J.R., P.B.R. Hazell, and M. Miranda, "New Approaches to Public/Private Crop Yield Insurance in Developing Countries.," Environment and Production Technology Division (EPTD) Discussion Paper 55, International Food Policy Research Institute (IFPRI) November 1999.

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

Tanaka, Y. and A. Munro, “Regional Variation in Risk and Time Preferences: Evidence from a Large-Scale Field Experiment in Rural Uganda,” *National Graduate Institute for Policy Studies Working Paper, Tokyo*, 2012.

Vyas, S. and L. Kumaranayake, “Constructing socio-economic status indices: how to use principal components analysis,” *Health Policy and Planning*, 2006, *21* (6), 459–468.

Wooldridge, J.M., *Econometric analysis of cross section and panel data*, The MIT press, 2002.

Yesuf, M. and R.A. Bluffstone, “Poverty, risk aversion, and path dependence in low-income countries: Experimental evidence from Ethiopia,” *American Journal of Agricultural Economics*, 2009, *91* (4), 1022–1037.

CONCLUSION

Summary. This thesis has analyzed the behaviour under risk and insurance purchase decisions of poor rural inhabitants using experimental data from Ethiopia and Brazil. It focused on answering empirical questions on the risk preferences and decision-making process under risk of individuals, as well as methodological questions regarding the suitability of the Multiple Price List (MPL) elicitation procedure in the developing country context. Chapter One estimated a wide range of models of behaviour under risk and showed that expected utility theory (EUT) does not provide a good overall description of the decisions made by participants in the Ethiopian experiment; instead, there is evidence of probability weighting and loss aversion. Chapter Two studied the determinants of risk preferences and showed that household wealth negatively affects both risk aversion and loss aversion exhibited in the decision problems of the Ethiopian experiment, but independent background risk has no significant impact on risk preferences. Further, there is evidence of narrow framing, as opposed to asset integration – participants make decisions in the experiment in isolation from outside wealth. Chapter Three analyzed data from an experiment conducted in Brazil that used the MPL elicitation procedure, which is more complex than the Ordered Lottery Selection method used in the Ethiopian experiment. It presented results which indicate that the MPL design enables a finer characterization of risk preferences and provides more power for the accurate estimation of the different decision models; however, there is evidence that subjects did not properly understand the decision problems and thus their choices do not reveal true risk preferences. Therefore, the complex MPL procedure may not be suitable for experiments conducted with poorly-educated subjects in developing countries, with simpler elicitation methods preferred in these settings. Chapter Four presented a theory outlining the relationship between rational demand for index insurance and wealth, and showed that the demand for index insurance in the Ethiopian experiment is hump-shaped – first increasing then decreasing – in wealth, in line with this theory. The results indicate that the low take-up of this product observed, particularly among the poorest (and most risk averse) individuals, in recent field studies may be due to rational choice rather than credit constraints or poor decision-making.

Future directions. Expected utility theory has long been the dominant theory of decision-making under risk in economics (Humphrey and Verschoor 2004). However, in line with recent studies on the subject, Chapter One showed that other decision models, which differ fundamentally from EUT, may better describe the behaviour under risk of rural inhabitants. Since risk preferences are a crucial input for both the effective design and evaluation of welfare-enhancing policies (Harrison et al. 2010), researchers and policymakers must ensure that the appropriate decision model and risk preferences – which accurately reflect the decision-making process – are identified and utilized.

The characterization of the decision-making under risk of poor individuals in developing countries remains incomplete, and the markedly different estimates of preference parameters obtained using the Ethiopian data (Chapter One) and the Brazilian data (Chapter Three) indicate that inhabitants of different regions may exhibit quantitatively and qualitatively different behaviour under risk. Thus, further research, using experimental data in particular, on accurately assessing the decision-making process under risk – as well as the corresponding preference parameters – in different regions is crucial. Most studies on the subject assume a particular decision-making process (generally EUT) *a priori* or analyze just one or two competing theories of choice – with the large number of possible decision models and the uncertainty regarding the decision-making process under risk of the rural poor, it is important to consider a wide range of models and let the data determine which theory (or theories) provides the best description of behaviour under risk. Further, assuming an incorrect decision-making process in analyses of behaviour under risk could yield severely misleading results. Additionally, given the growing popularity of reference-dependent decision models (such as cumulative prospect theory), and the importance of the reference point in these models, it is crucial to study what determines the reference point and whether it is "homegrown" or influenced by the framing of the experimental decision problems (a similar point is made by Abeler et al. 2011); this is important for increasing the economic applicability of reference-dependent decision theories.

Since experimental methods have emerged as an important avenue for eliciting risk attitudes, it is crucial to use an experimental design that accurately elicits the risk preferences of individuals in a particular region. Chapter Three showed that the complex Multiple Price List elicitation procedure, which has been used extensively in experiments conducted in developed countries, may not be properly understood by poorly-educated subjects in developing country settings. This highlights the importance of choosing the right experimental procedure (with the appropriate level of complexity) for a particular setting – one that is properly understood by participants and produces non-spurious responses that correctly reflect risk preferences – given that different experimental procedures may be appropriate for different settings and different subject pools (as noted by Charness and Viceisza 2012). However, this issue is often ignored by experimental economists, and has been given relatively little attention in the literature, particularly in the case of developing countries (Dave et al. 2010, Charness and Viceisza 2012). Therefore, further research on assessing the suitability of different experimental methods and determining the most appropriate procedure for experiments in developing countries is crucial. Additionally, it is important to analyze the tradeoffs between precision and comprehension for different elicitation methods – these are major dimensions along which experimental procedures should be assessed. A study on this topic should involve an experiment which includes decision problems using different elicitation procedures (such as the Multiple Price

List and Ordered Lottery Selection methods), conducted with the same sample of developing country participants (along the lines of Charness and Viceisza 2012) – this would enable a direct comparison of the effectiveness of different procedures for accurately eliciting risk preferences.[1]

As noted by Jacobson and Petrie (2009), most experimental studies ignore, or entirely discard, inconsistent choices, under the assumption that they represent uninformative noise. However, the results of Chapter Three highlight the importance of analyzing inconsistencies in choice (and the causes of these inconsistencies) for obtaining a better understanding of the decision-making process under risk of individuals. Furthermore, Jacobson and Petrie (2009), using data from Rwanda, find that inconsistencies in experimental choice are linked to sub-optimal financial and economic decisions in real life. Thus, given the potential importance of inconsistent choices, further research on the correlates and causes of these choices, as well as on the relationship between inconsistencies in experimental choice and economic decisions in real life, is vital.

Additionally, variations of the MPL design have been created which impose a single switch point and thus consistency in choice – these have been used, for example, by Andersen et al. (2006), Liu (2008) and Tanaka et al. (2010). However, using such variants of the MPL design could lead to a loss of important information on decision-making under risk in experiments. Additionally, subjects – particularly those with low levels of education and low cognitive ability – may face difficulty in understanding the complex MPL design, and thus make choices in a psuedo-random fashion that are consistent by design, but do not reflect or reveal true preferences over risk. If these responses are assumed to reflect true risk preferences, it would significantly bias the estimates of preference parameters and the results of these experimental studies could be severely misleading.[2] Thus, an important avenue for further research is to determine

[1] Further, Crosetto and Filippin (2013) argue that it is preferable to conduct a comparative study of this type using a between-subject, rather than a within-subject, analysis.

[2] Charness et al. (2013) also draw attention to the possibility that experiments using such elicitation procedures impose added assumptions – namely monotonicity in revealed preferences and transitivity in choice – which may not hold and could yield biased results.

how misleading the results of these studies (which use single-switch variations of the MPL procedure) may be if randomness in choice was commensurate to what is found in standard MPL experiments conducted in similar contexts. Additionally, it is essential to explore how such an experimental design could be improved by adding choices that help detect, and test for, inconsistencies in choice, errors in decision-making and evidence of poor comprehension.

Another promising avenue for further research is to investigate the possibility of creating variations of the MPL design that are easily understood, even by poorly-educated subjects. For example, using a MPL variant in which participants are presented choices in a visual, one-at-a-time format, Eckel et al. (2007) find a significant increase in comprehension and consistency. Thus, the Eckel et al. (2007) study shows that the MPL procedure can be implemented in a way that is understood by poorly-educated subjects in a developing country. Given the greater precision associated with the MPL procedure, improving the implementation and increasing the comprehension of the MPL method is preferred to abandoning it in favour of the Ordered Lottery Selection procedure, for experiments conducted in developing countries.

Alternatively, the possibility of increasing the precision of the simpler Ordered Lottery Selection design should also be explored. For example, an analogue of the Iterative MPL procedure (outlined by Andersen et al. 2006) could be considered for simpler Ordered Lottery Selection problems. Such a procedure would include a series of iterative Ordered Lottery Selection problems presented to subjects – the first would include options that cover a wide range of risk preferences; after a subject's choice in the first problem, the second problem presented would allow the subject to make a choice from refined options within the last option chosen (Andersen et al. 2006). Since the second problem contains options which only allow for risk preferences within the interval implied by the participant's first choice, this procedure enables a finer classification of risk attitudes as compared to a single Ordered Lottery Selection problem, while preserving the simplicity of the Ordered Lottery Selection design.

This process could be continued until the desired level of precision in elicited risk preferences is obtained. Additionally, an attractive alternative to both the MPL and Ordered Lottery Selection methods for experiments conducted in developing countries is the simple, single-choice investment task used by Gneezy and Potters (1997) to elicit risk preferences; this experimental task has already been widely implemented in developing countries (for example, Gneezy et al. 2009, Gong and Yang 2012).

The results of Chapter Two imply that richer households are less averse to both risk and losses than poorer households. This finding, combined with evidence from numerous studies on the path dependence of wealth accumulation, indicates that risk preferences may have a crucial impact on long-run poverty – poorer households are more risk and loss averse and thus may be more likely to undertake low-risk, low-return investment and production strategies, thereby hampering wealth accumulation and leading to persistent poverty (Mosley and Verschoor 2005, Yesuf and Bluffstone 2009). Further research on understanding the determinants of risk preferences is therefore crucial, as it would provide insights into the investment decisions of the rural poor and shed light on their economic circumstances and well-being (Fafchamps 2003, Harrison et al. 2010). Additionally, establishing the direction and magnitude of the impact of economic circumstances – such as wealth – on risk preferences is essential for a better understanding of decision-making under risk, and has crucial implications for the design of anti-poverty policies (Banerjee 2000). Falco (2012) notes that contributions to the debate on the endogeneity of risk attitudes and the direction of causality between economic factors and risk preferences represent important avenues for further research, with crucial implications for both economic policy and theoretical modeling.

There are very few studies on the determinants of risk attitudes that jointly consider both wealth and background risk, and even fewer that appropriately account for the endogeneity associated with both these variables. Lusk and Coble (2008) note that ignoring background risk could generate biased estimates and misleading inferences when analyzing risk-taking behaviour, and stress the importance of further experimental research assessing the impact of background risk on risk preferences. Therefore, future studies on this topic should jointly consider both wealth and background risk, as well as appropriately account for endogeneity, in order to accurately assess the determinants of risk preferences. In addition, further experimental research is required to analyze the non-experimental ("field") background risks experiment participants are likely to bring with them into an experiment, as the effect of these larger risks might dominate that of experimentally-induced background risk (Lusk and Coble 2008).

Tanaka et al. (2010) also note that while most studies analyzing the correlates of risk preferences utilize a single-parameter EUT framework, further research on this topic using non-EUT frameworks is crucial, given that evidence from numerous studies (including theirs) indicates that non-EUT decision models with multiple preference parameters better fit both experimental and field data than EUT models. In particular, it is important to analyze the determinants of loss aversion (in addition to risk aversion) – recent studies have found that loss aversion may be a more important characterization of the behaviour, and play a greater role in the decision-making, of poor villagers in developing countries than risk aversion (for example, Tanaka et al. 2010, Fafchamps 2009).

Development economists and insurance practitioners are rightly optimistic about the potential for index insurance to substantially increase welfare for many of the world's rural poor (Skees et al. 1999, Karlan and Morduch 2009). However, Chapter Four showed that the low take-up of weather indexed insurance observed in pilot projects, particularly among the poorest (and most risk averse) farmers, may be a result of rational choice due to a fundamental lack of desirability of the product, rather than barriers to the adoption of a new technology, such as lack of understanding, credit constraints, unwillingness to experiment or poor decision-making on the part of rural consumers. If demand for index insurance is "rationally" low for the most risk averse individuals – who are also likely to be the poorest – it implies that the measures outlined by Giné et al. (2008) and Cole et al. (2009) to address credit constraints and “behavioural” impediments to index insurance take-up may not be effective in increasing take-up by the poorest, who comprise vulnerable populations that would presumably benefit the most from insurance against weather risk. Additionally, improving understanding of the product (a strategy suggested by Lybbert et al. 2010) may not lead to significant increases in its take-up.

Thus, to increase index insurance purchase by the poorest individuals, other provisions (such as government-provided safety nets) may be required – in combination with index insurance contracts – to mitigate or protect against the negative basis risk associated with indexed products. Hill et al. (2011) note that the design of existing index insurance products is far from perfect, and research on this design – with a focus on reducing basis risk – is essential, given the large potential risk management and welfare benefits, as well as the rapidly expanding market for microinsurance. In particular, it would be informative to further explore the methods suggested by Dercon et al. (2011) and Hill and Nobles (2011) to promote index insurance take-up – they propose that marketing the products to pre-existing risk-sharing groups (such as iddirs in Ethiopia) and offering a range of flexible weather indexed insurance products, respectively, would substantially increase take-up.

However, Hill and Nobles (2011) also note that basis risk will continue to be an issue for any weather indexed insurance product, and further research is also required to determine how index insurance can be integrated with the other risk-coping mechanisms currently used by rural inhabitants, such as group-based risk sharing as well as informal credit and gift-giving. Alternatively, insurance providers need to design new insurance products – which involve little (or no) basis risk – that are attractive to very risk averse individuals.

Academic research on microinsurance – and particularly index insurance – is still in its infancy, and many crucial questions remain unanswered (Giné et al. 2008).

However, on their own, field experiments seem unlikely to advance our understanding of key questions such as how to design good insurance products or whether clients make good financial decisions (that is, optimal or rational decisions). To contribute to these questions we will also need normative theory and carefully designed lab experiments. Combined data from lab experiments and field surveys, in particular, provides fertile ground to further analyze decision-making under risk and insurance provision in developing economies.

References

Abeler, J., A. Falk, L. Götte, and D. Huffman, "Reference points and effort provision," *The American Economic Review*, 2011, *101* (2), 470–492.

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

Banerjee, A.V., "The Two Poverties," *Nordic Journal of Political Economy*, 2000, *26* (2), 129–141. Reprinted in Dercon, Stefan (ed), "Insurance Against Poverty", Oxford University Press, 2004.

Charness, G. and A. Viceisza, "Comprehension and risk elicitation in the field: Evidence from rural Senegal," *IFPRI Discussion Paper*, 2012.

Charness, Gary, Uri Gneezy, and Alex Imas, "Experiential methods: Eliciting risk preferences," *Journal of Economic Behavior & Organization*, 2013.

Cole, Shawn A., Xavier Giné, Jeremy B. Tobacman, Petia B. Topalova, Robert M. Townsend, and James I. Vickery, "Barriers to Household Risk Management: Evidence from India," Working Paper 09-116, Harvard Business School 2009.

Crosetto, Paolo and Antonio Filippin, "A Theoretical and Experimental Appraisal of Five Risk Elicitation Methods," Technical Report, Friedrich-Schiller-University Jena, Max Planck Institute of Economics 2013.

Dave, C., C.C. Eckel, C.A. Johnson, and C. Rojas, "Eliciting risk preferences: When is simple better?," *Journal of Risk and Uncertainty*, 2010, pp. 1–25.

Dercon, Stefan, Ruth Vargas Hill, Ingo Outes-Leon, A Bayrou, D Clarke, and A Seyoum-Taffesse, "Offering rainfall insurance to informal insurance groups: evidence from a field experiment in Ethiopia," Technical Report, Mimeo 2011.

Eckel, Catherine, Jim Engle-Warnick, and Cathleen Johnson, "Adaptive Elicitation of Risk Preferences," *University of Montreal Working Paper*, 2007.

Fafchamps, M., *Rural Poverty, Risk and Development*, Edward Elgar Publishing, 2003.

Fafchamps, Marcel, "Vulnerability, Risk Management, and Agricultural Development," *CSAE Working Paper Series*, 2009.

Falco, P., "Does risk matter for occupational choices? Experimental Evidence from an African labour market," *Manuscript submitted for publication*, 2012.

Giné, X., R. Townsend, and J. Vickery, "Patterns of rainfall insurance participation in rural India," *The World Bank Economic Review*, 2008, *22* (3), 539.

Gneezy, U. and J. Potters, "An experiment on risk taking and evaluation periods," *The Quarterly Journal of Economics*, 1997, *112* (2), 631–645.

Gneezy, Uri, Kenneth L Leonard, and John A List, "Gender differences in competition: Evidence from a matrilineal and a patriarchal society," *Econometrica*, 2009, *77* (5), 1637–1664.

Gong, Binglin and Chun-Lei Yang, "Gender differences in risk attitudes: Field experiments on the matrilineal Mosuo and the patriarchal Yi," *Journal of Economic Behavior & Organization*, 2012, *83* (1), 59–65.

Harrison, G.W., S.J. Humphrey, and A. Verschoor, "Choice under Uncertainty: Evidence from Ethiopia, India and Uganda," *The Economic Journal*, 2010, *120* (543), 80–104.

Hill, R. V. and M. Nobles, "Flexible insurance for heterogenous farmers: results from a small scale pilot in Ethiopia," *IFPRI discussion paper 01092*, 2011.

Hill, R.V., J. Hoddinott, and N. Kumar, "Adoption of weather index insurance: Learning from willingness to pay among a panel of households in rural Ethiopia," *IFPRI discussion paper 01088*, 2011.

Humphrey, S.J. and A. Verschoor, "Decision-making under risk among small farmers in East Uganda," *Journal of African Economies*, 2004, *13* (1), 44.

Jacobson, S. and R. Petrie, "Learning from mistakes: What do inconsistent choices over risk tell us?," *Journal of Risk and Uncertainty*, 2009, *38* (2), 143–158.

Karlan, D. and J. Morduch, "Access to Finance: Credit Markets, Insurance, and Saving," in Dani Rodrik and Mark Rosenzweig, eds., *Handbook of Development Economics*, Vol. 5, Amsterdam: North Holland, 2009, chapter 71.

Liu, Elaine M, "Time to change what to sow: Risk preferences and technology adoption decisions of cotton farmers in China," *Review of Economics and Statistics*, 2008.

Lusk, J.L. and K.H. Coble, "Risk aversion in the presence of background risk: Evidence from an economic experiment," *Research in Experimental Economics*, 2008, *12*, 315–340.

Lybbert, Travis J, Francisco B Galarza, John McPeak, Christopher B Barrett, Stephen R Boucher, Michael R Carter, Sommarat Chantarat, Aziz Fadlaoui, and Andrew Mude, "Dynamic Field Experiments in Development Economics: Risk Valuation in Morocco, Kenya, and Peru," *Agricultural and Resource Economics Review*, 2010, *39*, 1–17.

Mosley, P. and A. Verschoor, "Risk attitudes and the 'vicious circle of poverty'," *The European Journal of Development Research*, 2005, *17* (1), 59–88.

Skees, J.R., P.B.R. Hazell, and M. Miranda, "New Approaches to Public/Private Crop Yield Insurance in Developing Countries.," Environment and Production Technology Division (EPTD) Discussion Paper 55, International Food Policy Research Institute (IFPRI) November 1999.

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

Yesuf, M. and R.A. Bluffstone, "Poverty, risk aversion, and path dependence in low-income countries: Experimental evidence from Ethiopia," *American Journal of Agricultural Economics*, 2009, *91* (4), 1022–1037.