Index Insurance, Risk Preferences, and Deprivation in Low-Income Economies

Der Wirtschaftswissenschaftlichen Fakultät der
Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des akademischen Grades

Doktorin der Wirtschaftswissenschaften
- Dr. rer. pol. -

genehmigte Dissertation

von

Diplomvolkswirtin Veronika Bertram-Hümmer

geboren am 21. Mai 1980 in Lichtenfels

2015

Promotionsausschuss:

Referentin:	Prof. Dr. Susan Steiner
Koreferent:	Prof. Nathan Fiala (PhD)

Tag der Disputation: 25. November 2015

Contents

Abstract

Many people in developing countries live in extraordinarily risky environments with insufficient access to formal insurance systems and safety nets. Their decision-making under risk has crucial implications for their welfare and continued poverty. Therefore, achieving a better understanding of the behavior of poor populations under high risk exposure is an essential step for the effective design and evaluation of welfare-enhancing policies. The three main chapters of this dissertation, which are preceded by an introductory chapter, contribute to a better understanding of peoples' welfare and behavior under risk in a developing country context.

In chapter 2, we investigate which dimension of welfare best describes perceived deprivation. Based on theoretical considerations by Runciman (1966), it is empirically examined whether deprivation in visible or in invisible goods better explains reported feelings of economic deprivation by using a representative household survey in Kyrgyzstan. We select a range of visible consumption and asset items that reflect different welfare dimensions and create a visible wealth index by using principal component analysis (Kolenikov and Angeles 2004; 2009). The estimation results show that an index of visible wealth - which includes a comprehensive set of visible consumption and asset items - has significantly larger explanatory power on feelings of deprivation than a simple, commonly used income measure. The finding is robust to a variety of robustness and sensitivity checks. The chapter sheds light on the importance of visible goods for the identification of deprived population groups, which is needed for the design of policy programs intended to fight poverty and inequality.

In chapter 3, we utilize an artefactual field experiment in rural Ethiopia to investigate two methods to elicit risk preferences with low-educated individuals in developing countries. Choices from a simple Ordered Lottery Selection experiment introduced by Binswanger (1980; 1981) (OLS-BW) are compared with the choices from a more comprehensive Multiple Price List (MPL) format with repeated decisions between a safe amount and a lottery. The chapter investigates non-parametric measures as well as parametrically estimated risk preferences and noise based on Expected Utility Theory (EUT) and Rank-dependent Utility (RDU) theory. We find that both experiments reveal similar levels of risk preferences and moderate levels of noise when parametrically estimated. In contrast to the OLS-BW, the MPL allows us to identify moderate levels of inconsistent choices and to estimate more complex economic models revealing an inverse s-shaped probability weighting function. Due to the experimental design, we find the non-parametric OLS-BW measure becomes heavily distorted toward risk aversion. Our findings suggest that the OLS-BW method is sufficient when parametrically estimating measures to characterize risk preferences. They encourage the usage of the MPL procedure when relying merely on non-parametric methods, when directly analyzing inconsistent choices, and when analyzing risk preferences in different risk environments.

In chapter 4, we analyze the effects of index insurance payouts on households' recovery after disaster. Index insurance was introduced in order to enable agricultural households to better cope with losses caused by extreme weather. The chapter elaborates the effects of index insurance payouts after a major winter disaster occurred in Mongolia in 2009/10, exploiting three waves of a representative household survey in three provinces in western Mongolia. Post-disaster livestock recovery of insured, pastoralist households is compared with that of similar households without insurance. To control for self-selection into buying insurance, we use the bias-corrected matching estimator (Abadie and Imbens 2002; 2006; 2011) and exploit the phasing-in of the insurance. The results indicate that pastoralist households purchasing insurance before the shock recover faster from shock-induced asset losses than comparable non-insured households. We find a significant, positive and economically large effect of indemnity payments on herd size one and two years after the shock. In the medium term - three and four years after the shock - the effect slowly vanishes. An analysis of shock coping strategies as well as complementary qualitative interviews conducted in the field suggest that indemnity payments help herders to smooth their productive asset base and to relieve credit constraints. The chapter provides first evidence on the micro-level benefits of index insurance after a weather shock.

Keywords: deprivation; risk preferences; index insurance.

Zusammenfassung

Das Leben vieler Menschen in Entwicklungsländern ist in extremem Ausmaß von Risiken geprägt, gegen die sie kaum durch soziale Sicherungsnetze oder formale Versicherungen abgesichert sind. Das Wohlergehen der betroffenen Haushalte hängt dabei grundsätzlich von deren Fähigkeit ab, mit diesen Risiken umzugehen. Ein gutes Verständnis des Entscheidungsverhaltens armer Bevölkerungsgruppen unter Risiko ist daher eine Grundvoraussetzung für eine erfolgreiche Gestaltung und für die Evaluierung von Politikinitiativen zur Armutsbekämpfung. Diese Dissertation bildet einen Beitrag zu einem besseren Verständnis der Verhaltensweisen von Menschen in Entwicklungsländern, die mit einem hohen Ausmaß an Risiko umgehen müssen.

In Kapitel 2 wird untersucht, welche objektiven Wohlfahrtsindikatoren am besten individuell wahrgenommene Benachteiligung erklären. Basierend auf der Deprivationstheorie von Runciman (1966) testen wir, ob sichtbare oder nicht sichtbare Güter einen größeren Einfluss auf subjektiv wahrgenommene, ökonomische Benachteiligung haben. Dafür verwenden wir Daten einer repräsentativen Haushaltsumfrage in Kirgistan. Die Ergebnisse zeigen, dass ein Indikator für sichtbaren Wohlstand, der umfassend die sichtbaren Konsumgüter und Vermögensgegenstände eines Haushaltes reflektiert, einen signifikant größeren Einfluss auf gefühlte Deprivation hat, als das Einkommen eines Haushaltes welches üblicherweise als Deprivationsmaß in empirischen Studien benutzt wird. Eine Reihe an Robustheits- und Sensitivitätsanalysen bestätigen dieses Ergebnis. Das Kapitel zeigt die Relevanz, die der Sichtbarkeit von Gütern bei der Identifikation von benachteiligten Bevölkerungsgruppen zukommt, und somit deren Bedeutung für eine zielorientierte Ausrichtung von Programmen zur Bekämpfung von Armut und Ungleichheit.

In Kapitel 3 evaluieren wir anhand eines Experimentes mit Landwirten aus dem ländlichen Äthiopien zwei Methoden zur Messung von Risikopräferenzen von Menschen in Entwicklungsländern. Wir vergleichen das weitverbreitete *Ordered Lottery Selection* Experiment von Binswanger (1980; 1981) (OLS-BW) mit einem *Multiple Price List* (MPL) Format mit wiederholten Entscheidungen zwischen einer Lotterie und einem sicheren Betrag. In dem Kapitel werden die ermittelten Risikopräferenzen und das

Ausmaß an Störfehlern sowohl nicht-parametrisch als auch parametrisch analysiert. Die Ergebnisse zeigen, dass beide Methoden nach einer parametrischen Schätzung zu ähnlichen Ergebnissen der Risikopräferenzen und Störfehler kommen. Im Vergleich zum einfachen OLS-BW Format ermöglicht das MPL jedoch eine genauere Charakterisierung von Risikopräferenzen und zeigt unterschiedliche Dimensionen von Risikopräferenzen auf. Die empirische Analyse zeigt zudem, dass die direkten, nicht-parametrischen Maße aus dem OLS-BW Experiment stark verzerrt sind. Um Fehlinterpretationen auszuschließen sollten Risikopräferenzen aus dem OLS-BW Experiment daher parametrisch geschätzt werden. Die Ergebnisse zeigen, dass das OLS-BW Experiment ausreichend ist, um ein einfaches Maß für Risikopräferenzen zu schätzen. Für eine genaue Ermittlung von Risikopräferenzen, sowie für eine direkte Identifikation von inkonsistenten Entscheidungen innerhalb des Experimentes, ist jedoch das MPL Experiment zu empfehlen.

Kapitel 4 untersucht die Auswirkungen von Auszahlungen einer index-basierten Versicherung nach einer Wetterkatastrophe. Index-basierte Versicherungen ermöglichen landwirtschaftlichen Haushalten sich gegen Verluste durch extreme Wetterereignisse zu versichern. Das Kapitel analysiert die Effekte von Auszahlungen nach dem katastrophalen mongolischen Winter 2009/2010 anhand von drei Datenwellen einer repräsentativen Haushaltsumfrage in der Westmongolei. Wir vergleichen die Entwicklung des Viehbestandes von versicherten Nomadenhaushalten mit der von vergleichbaren, nicht-versicherten Haushalten. Um eine mögliche Verzerrung auszuschließen, die sich dadurch ergeben könnte, dass sich Haushalte mit bestimmten Charakteristiken dazu entscheiden die index-basierte Versicherung zu kaufen, verwenden wir den *Bias-Corrected Matching Estimator* (Abadie and Imbens 2002; 2006; 2011), und nutzen die schrittweise Einführung des Versicherungsprogramms in unterschiedlichen Regionen aus. Die Schätzergebnisse deuten darauf hin, dass versicherte Haushalte besser mit den Verlusten umgehen können. Sie zeigen, dass die Auszahlungen aus der index-basierten Versicherung einen signifikant positiven Effekt auf den Viehbestand ein und zwei Jahre nach der Katastrophe haben. Mittelfristig geht dieser Effekt teilweise wieder zurück. Zur genaueren Erklärung haben wir die angewandten Risikomanagementstrategien quantitativ analysiert, sowie zusätzlich qualitative Interviews im Feld geführt. Die Auswertung dieser Daten weist darauf hin, dass versicherte Haushalte bereits während der Katastrophe einfacher an Liquidität durch Bankkredite kommen konnten. Des Weiteren waren versicherte Haushalte weitaus weniger gezwungen, eigenes Vieh nach der Katastrophe zu konsumieren oder zu verkaufen um die Grundlagen des täglichen Bedarfs sicherzustellen. Als eine der ersten empirischen Arbeiten zu den Auswirkungen von index-basierten Versicherungen zeigt das Kapitel die positiven Effekte von Versicherungsauszahlungen nach einer Katastrophe auf Haushaltsebene auf.

Schlagworte: Deprivation; Risikopräferenzen; Index-Versicherung.

Nomenclature

AIPW	Augmented inverse-probability weighting
ATE	Average treatment effect
BIP	Base insurance product
cara	Constant absolute risk aversion
CE	Certain equivalent
crra	Constant relative risk aversion
DRP	Disaster risk product
drra	Decreasing relative risk aversion
ep	Expo-power
ETB	Ethiopian Birr
EUT	Expected utility theory
EV	Expected value
FAO	Food and Agriculture Organization of the United Nations
FoSD	First-order stochastic dominance
IBLI	Index-based Livestock Insurance
IFRC	International Federation of Red Cross and Red Crescent Societies
IPWRA	Inverse probability-weighted regression adjustment
LIK	Life in Kyrgyzstan survey
LRI	Livestock risk insurance
MNT	Mongolian Tugrik
MPL	Multiple price list

MRCS	Mongolian Red Cross Society
NEMA	National Emergency Management Agency of Mongolia
NSO	National Statistical Office Mongolia
OLS	Ordinary least squares
OLS-BW	Ordered lottery selection of Binswanger
PCA	Principal component analysis
PIU	Project implementation unit
PSU	Primary sampling unit
RDU	Rank-dependent utility theory
SoSD	Second-order stochastic dominance
SUTVA	Stable unit treatment value assumption
UNDP	United Nations Development Program
US$	US Dollar

Chapter 1

Introduction

1.1 Motivation

Many people in developing countries face extraordinarily risky environments. They are exposed to a multitude of shocks, such as health shocks, weather shocks and asset shocks, which lead to large fluctuations in their income and consumption. At the same time, the lack of formal insurance and social safety net systems, as well as financial institutions increases their vulnerability. Fafchamps (2003, p. 196) describes their situation as follows:

> *Perhaps the only way to describe it to people who have never been there is to compare it to a war economy: death strikes at random a large proportion of the population, especially children; the provision of health services is either non-existent or insufficient; trade with the rest of the world is difficult so that many commodities are rationed or unavailable and local prices are erratic; food is at times very scarce; and steady wage employment is non-existent so that people must make a living from self-employment in little jobs.*
> *To deal with such a harsh environment, people are equipped with very little in terms of advanced technology and accumulated assets. Financial institutions are either absent or inefficient and expensive, and in many places, inflation is rife so that the cost of hoarding money is high.*

Given the large risk exposure and the limited institutional support, an individual's ability to cope with risks determines to a large extent his or her level of welfare and poverty. In particular, the poor are often not able to adequately manage risks. Uninsured risk and high levels of risk aversion cause households to withhold investments or invest in low-risk, low-return technologies. For example, poor agricultural households decide to grow low-return, but drought-resistant, subsistence crops (Dercon 2005) or they refrain from using fertilizer (Dercon and Christiaensen 2011). This kind of sub-optimal risk preparation implies missed opportunities for future welfare. When a shock occurs, poorer households are often forced to fall back to detrimental coping strategies, such as reduced nutrition (Dasgupta 1997) or distress sales of productive assets (Zimmerman and Carter 2003; Carter and Barrett 2006). In this way, the effects of shocks manifest themselves not only in terms of short-term negative impacts, but also in terms of sustained poverty and a reduced ability to deal with future shocks. The poor are trapped into poverty by sub-optimal risk management decisions that further raise their vulnerability to adverse shocks (Fafchamps 2003).

Understanding behavior under risk is crucial when designing and developing effective policy instruments that target the poor and vulnerable. Governments and international organizations wanting to support individuals' in their risk management, for instance through the implementation of micro-level insurance schemes, need accurate information on risk exposure, welfare and risk preferences of the target group when designing and developing a program (Harrison et al. 2010). This thesis investigates welfare and behavior of people characterized by high levels of vulnerability and risk exposure and is devoted to a better understanding of individual behavior under risk to improve the design and the evaluation of risk-reducing policies that target the poor with the overarching aim of raising welfare in developing countries.

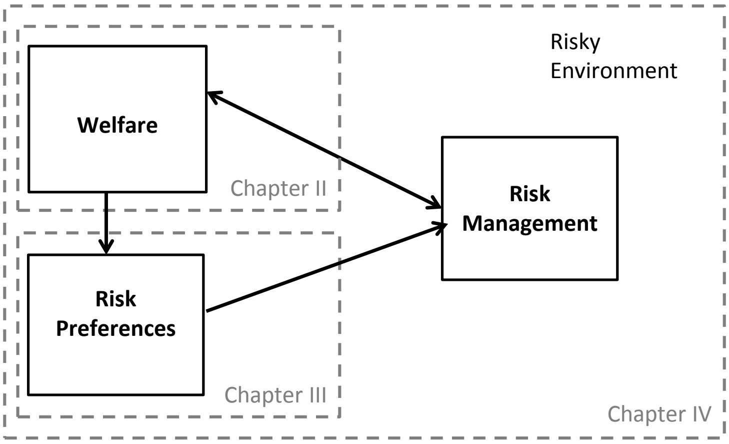
1.2 Background

Individual risk preferences and risk management behavior, as well as the level of welfare and poverty are interlinked in various ways (see illustration in figure 1.1):

– Welfare levels and risk management behavior are closely interrelated. On the one hand, the level of welfare determines an individual's capability to manage risk and it directly affects risk management decisions. To adequately prepare for shocks, it is necessary to have sufficient liquidity or assets to invest in risk protection or insurance. Furthermore, the level of wealth determines the capabilities of an individual to smooth consumption, income and productive assets in the aftermath of an adverse shock. On the other hand, the way an individual manages risks has a direct influence on income and consumption levels and, consequently, on the level of welfare in the short and long run. For example, sub-optimal risk management decisions, such as the sale of productive assets, instantly reduce income and consumption, leading to a lower level of welfare in the future (Dercon 2005).

– Risk preferences also have a heavy influence on individual risk management decisions and, thus, indirectly on the level of welfare and poverty. For example, it is

Figure 1.1: Thematic Overview of the Dissertation



Note: The figure illustrates the relationships between individuals' welfare, risk preferences and risk management decisions within a background of high risk exposure. The arrows reflect causal links. The grey-dashed squares frame the topics covered by the respective chapters of the dissertation. The figure is inspired by Liebenehm (2015, p. 19).

found that risk averse individuals are more hesitant to invest in new technologies that involve higher risks, but would generate higher profits and welfare (Liu 2013). Furthermore, risk attitudes influence insurance decisions, such as the uptake of formal insurance (for instance Giné et al. 2007; Giesbert et al. 2011) or joining an informal risk sharing group (Attanasio et al. 2012), which affect their ability to smooth income and consumption in the case of a shock. Risk preferences itself are influenced by welfare (e.g. Liebenehm and Waibel 2014) and welfare shocks (e.g. Menkhoff and Sakha 2014b; Callen et al. 2014) which again impacts the capabilities to manage risks.

The main chapters of this thesis focus on three particular aspects within the background of individual welfare, preferences and decisions in a risky environment:

Subjective deprivation In the conventional economic literature, welfare is usually analyzed by simple measures of income, consumption or assets (for instance Kuznets 1955; Lerman and Yitzhaki 1985; Birdsall 1997). However, there is ample evidence from different academic disciplines that it is the subjective perception of the own situation - which can only to a certain degree be explained by such objective measures - that drives behavior (e.g. Diener 2009; Shah 2012). One important factor why subjective perception differs from the objective and measurable reality is that people compare themselves with others (Ariely 2008). Perceived welfare is lower than objective welfare once an individual feels more deprived (or worse-off) than others in his or her reference group (Runciman 1966; Easterlin 1995). It is found that relative deprivation heavily affects behavior and decisions, such as consumption (Kuhn 2011; Linssen et al. 2011), or taking health risks (Deaton 2001; Eibner and Evans 2005).

Standard economic literature does not adequately address questions related to subjective welfare and to the constituents of relative deprivation. Most empirical studies use objective measures that do not satisfy the theoretical foundations of deprivation which would require the object of comparison to be visible (Runciman 1966). They only give various arbitrary assignments to both the group of reference and the choice of the object of comparison (Luttmer 2005; Ferrer-i-Carbonell 2005; Kingdon and Knight 2007; Klasen 2000, among others). Hence, a better understanding of subjective welfare and relative deprivation is required for enhancements in the economic analysis of individual behavior and decision-making.

Eliciting risk preferences To account for risk preferences in economic analysis and policy-making, it is necessary to adequately elicit individual risk preferences. This is not a trivial task. Unlike quantifiable information, such as a person's income or consumption, measuring risk preferences is not straightforward as it reflects a subjective notion that is considered to be an underlying trait (Dohmen et al. 2011) that might differ depending on the situation, for instance the area of risk-taking (Weber et al. 2002; Hanoch et al. 2006; Anderson and Mellor 2009) or the riskiness of the environment (Harrison and Rutström 2008; Harbaugh et al. 2010).

There is no established method of how to adequately elicit and measure risk preferences, in particular with low-educated individuals in developing countries (Charness et al. 2013). Many studies use simple elicitation measures, for example derived from willingness-to-take-risk questions, that are easily understood, but very limited in their information value on different dimensions of risk preferences, such as subjective probability weighting (Harrison and Rutström 2008). Studies that use more complex measures face the challenge that the elicitation task might not be well-understood by low-educated individuals, which consequently leads to noisy risk preference measures (Dave et al. 2010). In a review, Charness and Viczeisca (2013) point out that there are few studies that evaluate different elicitation methods based on parametric methods using experiments conducted in developing countries. For example, Menkhoff and Sakha (2014a) investigate different non-parametric risk preference measures with a sample in Thailand. To the best of our knowledge, there is no study that parametrically investigates different risk preference elicitation methods suitable for low-educated samples, which is necessary to improve experimental design and to enhance the measurement of risk preferences in developing countries.

Index insurance as a risk management tool Index insurance offers agricultural households to formally insure against losses due to weather disasters and thus might be an effective policy tool to mitigate excessive risk aversion and its negative welfare impacts. Unlike traditional insurance, index insurance payouts depend on an aggregated index reflecting the weather conditions in the region (Skees and Barnett 2006). This has several advantages, such as decreased moral hazard, adverse selection and transaction costs. These features underly the high expectations of practitioners and researchers. Since the late 1990s, index insurance programs have been increasingly implemented to enable agricultural households to better cope with losses caused by weather disasters (Carter et al. 2014). The academic literature on index insurance is still scarce. Empirical studies focus on the *ex-ante* effects of index insurance uptake (for instance Mobarak and Rosenzweig 2013; Hill and Viczeisca 2012). These studies find that index insurance causes households to invest in high-profit, high-risk investments leading to improved levels of welfare. There is very limited knowledge on the *ex-post* effects of index insurance payouts after disaster (Carter et al. 2014). Only one study of Janzen and Carter (2013) analyzes the *ex-post* effects of payouts from an index insurance pilot program in Kenya. They find positive effects on the households' expected coping strategies in the aftermath of the disaster. To evaluate whether index insurance schemes are a sustainable solution for agricultural households, evidence on the effects of index insurance payouts is needed.

1.3 Outline

The main part of this dissertation consists of three chapters, each a distinct research paper. The chapters investigate three different aspects of the relationship between individual welfare, risk preferences, and risk management behavior using household survey data from Kyrgyzstan and western Mongolia, and from an artefactual experiment in rural Ethiopia. The thesis makes two methodological contributions - finding an adequate proxy measure for subjective deprivation and identifying a procedure to accurately elicit risk preferences in a rural low-income economy. It further makes an empirical contribution on the micro-level impacts of index insurance payouts after a disaster on households' welfare. Thus, it contributes to the literature on development economics, microeconomics and experimental economics. In the following, I give a brief outline of each chapter.

Chapter 2 investigates relative deprivation, exploiting information on individually perceived deprivation. The chapter elaborates which of the objective deprivation measures commonly used in empirical studies serves as the best proxy for reported feelings of deprivation. We investigate different objective measures of welfare. Based on theoretical considerations, we test whether visible deprivation has a stronger explanatory power to determine perceived deprivation than income as a form of invisible deprivation (Runciman 1966).

The empirical analysis relies on the 2011 wave of the *Life in Kyrgyzstan Survey* (LIK), a large household data set that includes about 3,000 households. This data set is particularly suited to examine deprivation as it includes unique information on an individual's perceived level of relative deprivation with a clearly specified comparison group, the "other people in the town or village". The explicit information of the relevant comparison group allows us to steer away from arbitrary assumptions regarding the group of comparison (for instance all households in a region or country) as usually done in the empirical literature. The information on an individual's perceived level of deprivation is used as the main dependent variable and our benchmark for the comparison between visible and invisible deprivation. Furthermore, the LIK survey includes detailed information on income and on a wide range of consumption and asset items that serve as the components of the different indicators of objective deprivation. In particular, we create a visible wealth index by identifying and selecting a range of visible consumption and asset items that reflect different welfare dimensions (housing, transportation, livestock, durables and consumption). We then generate corresponding weights for each dimension separately using principal component analysis (PCA) (Kolenikov and Angeles 2004; 2009) and aggregate them into a unitary composite index of deprivation.

Our testing strategy is designed to compare the influence between visible and non-visible deprivation indicators on the levels of perceived deprivation. We evaluate the performance of deprivation in visible wealth against (non-visible) income to determine the driving elements behind individually perceived relative deprivation. We estimate the separate and joint effects of relative deprivation in income and visible wealth on perceived deprivation with ordinary least squares regressions and run a number of specifications of the model, including several robustness and sensitivity checks.

The results show that deprivation in visible wealth has a significantly stronger effect than deprivation in income in determining levels of perceived relative deprivation. The finding highlights the importance of visibility for an object of comparison and provides empirical evidence on the theoretical constituent of visibility for deprivation. It proposes the usage of visible goods for deprivation measures in empirical studies. The chapter contributes to a better understanding of the determinants of subjective deprivation and welfare of individuals in developing countries. It is based on joint work with Ghassan Baliki and was published in *Social Indicator Research* in December, 2014.

Chapter 3 evaluates two experimental approaches that are commonly used to elicit risk preferences in developing countries, the simple Ordered Lottery Selection of Binswanger (1980; 1981) (OLS-BW) and a more comprehensive Multiple Price List (MPL) format with multiple choices between a lottery and a safe amount.

We elicited individual risk preferences within an artefactual field experiment with 875 farmers in the Tigray region of Ethiopia. In the OLS-BW experiment, subjects make a choice between six lotteries reflecting different levels of risk aversion. The lotteries have a fixed winning probability of 50% and differ in the size of their outcomes. Each choice corresponds to a certain level of risk aversion, ranging from risk aversion to risk neutrality. As subjects are required to make only a *single* choice between the six lotteries, which do not differ in terms of their probability, the OLS-BW lottery can be easily understood by low-educated subjects; hence the OLS-BW is in general the default choice among experimental researchers who want to elicit risk preferences in a developing country. In the MPL experiment, subjects repeatedly choose between a lottery and an increasingly safe amount. By requesting multiple decisions over six price lists with different risk environments, including a low, middle and high risk environment, as well as winning and losing frames, the MPL method elicits more comprehensive information on individual risk preferences. The MPL procedure is increasingly used in studies investigating risk preferences in developing countries (e.g. Henrich and McElrath 2002; Callen et al. 2014; Vieider et al. 2015).

We evaluate the two risk elicitation methods by analyzing both non-parametric and parametric measures of risk preferences and their ability to reflect risk preferences in different risk environments. This comprises an investigation of estimated stochastic noise and non-parametric noise, as measured by violations of stochastic dominance. Our estimations of risk preference measures and noise are based on Expected Utility Theory (EUT) and Rank-dependent Utility (RDU) theory using maximum likelihood techniques. This chapter contributes methodologically to the literature on risk preferences. To the best of our knowledge, it is the first study that parametrically evaluates risk preference measures based on elicitation methods tailored for low-educated samples in developing countries. Our study exploits risk choices from the sample of Ethiopian farmers in both the OLS-BW and the MPL experiment. In contrast to studies that evaluate merely non-parametric measures of different methods (e.g. Loomes and Pogrebna 2005), our analysis builds on the parametric estimation of structural decision-making models. This allows us to disentangle risk preferences from stochastic noise and to investigate the interactions between risk preferences and wealth, as well as the level of risk exposure. In comparison to studies based on hypothetical tasks (e.g. De Brauw and Eozenou 2014), our analysis relies on incentivized tasks that are found to generate more reliable risk preference measures (Kachelmeier and Shehata 1992). Finally, the experimental approach allows us to isolate risk preferences within a laboratory environment, which is not feasible with non-experimental methods where the elicited risk preferences might only be valid within a certain context (Harrison and Rutström 2008).

In general, we find that both measures reveal similar levels of risk preferences with moderate levels of stochastic noise when parametrically estimated. The estimated risk preference measures characterize our sample of Ethiopian farmers as risk loving. We further find decreasing relative risk aversion (*drra* preferences), which implies risk lovingness is increasing relative to wealth. In contrast to the design of the OLS-BW experiment, the MPL allows for the detection of moderate levels of inconsistent choices from our sample and to estimate more complex RDU models testing for subjective probability weighting. We find that subjects in our sample overweight low probability risks and underweight high probability risks. Due to the cap at risk neutrality in the OLS-BW design, we find the non-parametric OLS-BW risk preference measure becomes heavily distorted toward risk aversion. When investigating the drivers of stochastic noise, we find noise in the MPL choices is significantly higher when the game is played in the afternoon (rather than the morning) sessions and that the enumerators significantly influence levels of noise in both experiments. Our findings suggest that the simple OLS-BW experiment is sufficient when parametrically estimating risk preferences, while the MPL experiment is preferable when relying on non-parametric methods to analyze risk preferences, when investigating inconsistent choices, and when estimating more complex economic models to analyze risk preferences in different risk environments. The chapter contributes to improvements in the elicitation design and the measurement of risk preferences in developing countries, and consequently to enhancements in the analysis of risk management behavior. It is based on joint work with Karlijn Morsink and is currently being prepared for submission to the *Journal of Development Economics*.

Chapter 4 elaborates the effects of index insurance payouts after disaster on households' recovery. In particular, we investigate the impacts of index insurance payouts after a catastrophic, once-in-50-years winter disaster in Mongolia on pastoralist households' asset recovery.

The chapter builds on three waves of the *Coping with Shocks in Mongolia Household Panel Survey*, which is implemented by the German Institute for Economic Research (DIW Berlin) in three provinces in western Mongolia. The survey includes detailed information on the households' socio-economic situation, on their risk management and on index insurance uptake and payouts. The key outcome variables in the analysis are the households' livestock holdings at four points in time (2011, 2012, 2013 and 2014) after the 2009/10 winter disaster. The treatment variable is a binary variable indicating whether a household was insured and received payouts after the 2009/10 winter. Our sample includes 59 insured herding households and 583 households that are comparable to the insured households in terms of their disaster experience.

The main challenge when evaluating the impacts of a commercial product, such as IBLI, is that it is - by its nature - freely available and households voluntarily decide for or against buying IBLI. To control for selection into buying insurance before the catastrophic winter and receiving insurance payouts in 2010, we use the bias-corrected matching estimator (Abadie and Imbens 2002; 2006; 2011). Thereby, we utilize the availability of a large number of retrospective pre-disaster household characteristics in our survey, which we use to predict insurance uptake. Furthermore, we exploit the phasing-in of the IBLI scheme. When the 2009/10 winter disaster occurred, IBLI was still in its pilot stage and only available in one of the three survey provinces. Therefore, we compare post-disaster livestock recovery of insured households with non-insured households in regions where insurance was not yet available. We further investigate the channels of the effect based on an analysis of the shock coping strategies and complementary qualitative interviews conducted in the field. The estimation results rest on the assumptions that there are no spillover effects of insurance payouts and any remaining unobservable factors explaining insurance uptake.

We find a significant, positive and economically large effect of index insurance payments on herd size one and two years after the shock. In the medium term – three and four years after the shock – the effect slowly vanishes. Results are robust to defining post-shock livestock recovery in different ways, varying the number of matches per observation, the choice of covariates, and the use of alternative propensity score estimators. The analysis of shock coping strategies and the information from qualitative interviews in the field suggest that indemnity payments help herders to avoid selling and slaughtering animals, and as such to smooth their productive asset base. Also, index insurance appears to have relieved households from credit constraints. The chapter contributes to the literature on the impacts of index insurance and provides first evidence on the micro-level benefits of index insurance payouts after a weather shock. It is based on joint work with Kati Krähnert and is currently being prepared for submission to *World Development*.

Chapter 2

The Role of Visible Wealth for Deprivation*

Veronika Bertram-Hümmer
Ghassan Baliki

*This chapter is published in *Social Indicators Research*, Dec 2015, Vol. 124, Issue 3, pp 765-783, http://dx.doi.org/10.1007/s11205-014-0824-2.

2.1 Introduction

The notion that happiness and well-being are not just only dependent on an individual's own economic standing, but also on their relationship to others has been strongly established within the social science literature (Easterlin 1995; 2001; Mc Bride 2010; Alpizar 2005; Kingdon and Knight 2007; D'Ambrosio and Frick 2007; among others). Researchers have voluminously examined the effects of relative deprivation on economic behavior and decision-making. For example, it has been shown that feelings of relative deprivation have direct repercussions on every day decisions, such as consumption (Kuhn 2011; Linssen 2011) and taking health risks (Eibner and Evans 2005; Deaton 2001). Moreover, relative deprivation is found to drive more potent decisions in life, such as the use of violence (Moghaddam 2005; Macours 2011), migration (Stark 1984; 1991, Bhandari 2004) and education (Ferguson and Michaelson 2013). At the same time, this vast literature has not adequately addressed questions related to the constituents of relative deprivation. These constituents include mainly the "object of comparison" and the "reference group". Most empirical studies on relative deprivation do not address these issues explicitly, and only give various arbitrary assignments to both the group of reference and the object of comparison (Luttmer 2005; Ferrer-i-Carbonell 2005; Kingdon and Knight 2007; Klasen 2000; among others).

The term relative deprivation was initially coined by Runciman (1966), where a person A is considered relatively deprived of an object X when (1) A does not have X and wants it, and sees it as feasible to have it, and (2) sees some other person or persons G with X.[1] (1) and (2) reflect the deprivation and the relativity criteria respectively. Hence, by definition, in order for both criteria to be satisfied, an essential condition is required: object X and a reference group G must be seen by A. In the following, we call this the *visibility condition* for relative deprivation.

The main choice of object X in empirical studies is income since it is widely available and it is considered to be proportional to other dimensions of economic well-being (e.g. Deaton 2001; Kingdon and Knight 2007; D'Ambrosio and Frick 2012). But does income satisfy the full conditions to be used as a measure of relative deprivation? Despite possible inaccuracies in the measurement of income due to the underreporting by the rich and the miscalculation by the self-employed and the poor (Biemer et al. 2011; Van Praag et al. 1983), income may still qualify for the assessment of *absolute* deprivation since it is privately known. Yet for the calculation of *relative* deprivation measures, income

[1] The term "deprivation" is used equivalently to "relative deprivation" throughout the rest of the paper.

may not be as valuable since the information on earnings of all the comparison group members must be available a priori to all individuals. Realistically, this information is difficult to obtain given that income is neither easily observable nor necessarily visible. Moreover, income and total earnings neglect the disaggregation of a household's decision into expenditures and savings which is vital for understanding different mechanisms of relative deprivation. In order to overcome the limitations of income measures, other studies use consumption, positional goods or assets values as objects of comparison, which may comprise better indices in analyzing relative deprivation (e.g. Klasen 2000; Fafchamps and Shilpi 2008; Ravallion and Loshkin 2008; Linssen 2011). However, these studies rely heavily on aggregates without differentiating between the observable and unobservable elements of their indices and hence do not account directly for the visibility condition.

Another shortcoming in the empirical literature is the identification of the comparison group G. Few studies explicitly ask respondents with whom they are comparing (Knight et al. 2009; Clark and Senik 2010). Knight et al. (2009) find that people in the immediate vicinity - namely the neighbors and villagers - are the most frequently chosen comparators in rural China. Clark and Senik (2010), on the other hand, find that colleagues are the most mentioned group of reference when comparing income in Europe. Without available information on the group of comparison, a common default choice is "all citizens of a country or a region" (e.g. Easterlin 1995; Deaton 2010; Klasen 2000; Grimm et al. 2002; Bhandari 2004). This is generally based on strong assumptions given that it is simply hard to believe that a farmer in the countryside would compare himself with a businessman in the city whom he most probably has never seen.

To the best of our knowledge, no study yet has used an index to measure relative deprivation, consisting of visible items, and only few determine the reference group. Moreover, and to our surprise, we could not find any study that compares and assesses the effectiveness and reliability of various indices for measuring relative deprivation. Motivated by Runciman's pioneering theory and the empirical gap in the analysis of the relative deprivation indices, this paper has two aims: (i) to introduce a measure of visible wealth as an object for relative deprivation which is constructed to meet the visibility condition as precisely as possible with household data, and (ii) to evaluate its performance against income to determine the driving elements behind true relative deprivation.

In order to accomplish those two aims, we use data from a socio-economic household survey in Kyrgyzstan in 2011. This data set is of special significance because it includes unique information on an individual's perceived level of relative deprivation. Perceived relative deprivation serves as our proxy for the true level of relative deprivation and will be the main dependent variable of analysis. Moreover, the question on perceived relative deprivation clearly specifies the comparison group for the respondents as the other people in the town or village. This advantage allows us to steer away from any assumptions regarding the assignment of a relevant comparison group. Further, it facilitates matching the aggregation level of the calculated deprivation measures with reported deprivation, given that we can easily generate the indices for both income and visible wealth at the town and village levels too. For the creation of the visible wealth index, we identify and select a range of visible consumption and asset items in different wealth dimensions (housing, transportation, livestock, durables and consumption). We generate corresponding weights for each dimension separately using principal component analysis (PCA) and then aggregate them into a unitary composite index. By testing the separate and joint effects of the relative deprivation indices in income and visible wealth, we find that deprivation in visible wealth has a significantly stronger effect than income in determining levels of perceived relative deprivation. We run several specifications of the model and discuss a number of robustness and sensitivity checks. Our finding sheds light on the importance of visibility in relative deprivation measures and urges future research to not fully rely on income when analyzing the effects of relative deprivation on well-being and behavior.

The following section introduces the testing strategy and explains our empirical model. Section 2.3 explains the data and how the visible wealth index and the actual deprivation measures are calculated. In section 2.4, descriptive and estimation results are shown, followed by robustness checks. Section 2.5 concludes.

2.2 Testing Strategy

Building on Runciman's work (1966), our testing strategy is designed to compare the influence between observable and non-observable relative deprivation indicators on the levels of perceived deprivation. In order to establish this structure, we first assume that the utility of every individual depends only on their economic deprivation position compared to others, without the inclusion of the absolute levels. Hence, let the utility function of an individual i be standardized and depend negatively on the level of relative deprivation D^*_{ic}, then

$$U_i = 1 - D^*_{ic}, \tag{2.1}$$

with $D^*_{ic} \in [0,1]$ being the normalized measure of the true feeling of relative deprivation of individual i in comparison group c. This simply implies that the most (least) deprived individual within the comparison group has a utility equal to zero (one). It is important to note that here D^*_{ic} does not reflect the actual levels of relative deprivation, but rather the feeling of relative deprivation.

In order to make a clear distinction between true feelings of relative deprivation D^*_{ic} and actual levels of relative deprivation D^a_{ic}, imagine a group G with three individuals (i_1, i_2, and i_3). All three individuals are fully informative, rational, and have identical psychological and demographic traits and characteristics. i_1, i_2, and i_3 own 1, 2, and 3 cows respectively. Let their utility only depend on the number of cows they own relative to the others. Then, the actual relative deprivation in cows for i_1 for instance, is 2 compared to i_3 and 1 compared to i_2, for i_2 it is 1 compared to i_3, while the actual relative deprivation for i_3 is 0. Given that all individuals are identical, rational, and fully informative, then their true feelings of relative deprivation are equal to the actual deprivation. In other words, i_1 knows he is the most deprived and feels this way, while i_3 knows he is the least deprived (or not deprived) and feels this way. However, the true feelings of relative deprivation under any other assumptions would not necessary be equal to the actual levels of relative deprivation. For example in the case without full information, i_2 might think that i_1 is secretly hiding another two cows in the barn (which is not the case). Hence, i_2 feels the most deprived although in reality he or she is not, and therefore the utility will be equal to 0. Thus, the true feeling of deprivation D^*_{ic} might not necessarily coincide with the actual deprivation D^a_{ic}. The relationship can be shown as follows:

$$D^*_{ic} = \beta D^a_{ic} + \sum_k^K \gamma_k x_{ki} + \sum_l^L \delta_l w_{lc}, \tag{2.2}$$

where x_{ik} is the set of K individual-specific factors and w_{lc} the set of L comparison group-specific factors that may influence the sense of relative deprivation independently from the economic elements. These can include individual characteristics (age, gender, etc.), as well as group characteristics such as common values and norms within the comparison group. In order to account for the visibility condition in the computation of actual relative deprivation, let D^a_{ic} be mainly comprised of two mutually exclusive constituents, the visible component D^v_{ic} and the invisible component D^n_{ic}:

$$D^a_{ic} = \theta_1 D^v_{ic} + \theta_2 D^n_{ic}, \tag{2.3}$$

where θ_1 and θ_1 signify the weights that each individual assigns to the two components in assessing the relative position to others, such as $0 < \theta_1 + \theta_2 \leq 1$. We aim to test the magnitude of θ_1 and θ_2 in order to be able to assess the role of visibility in determining the factors in play behind true feelings of relative deprivation.

Using empirical data imposes two challenges: neither the true feelings of relative deprivation D^*_{ic} can be directly measured, nor the explicit differentiation between visible and invisible items can be clear-cut observed. Therefore, in order to overcome those two challenges, we need to find proxies for D^*_{ic}, D^v_{ic} and D^n_{ic}. First, in order to capture the closest representation of the true feelings of relative deprivation, we use reported levels of perceived relative deprivation D_{ic} as a proxy of the true feeling of deprivation, with

$$D_{ic} = \tau \iff \kappa_\tau \leq D^*_{ic} < \kappa_{\tau+1}. \tag{2.4}$$

Reported perceived relative deprivation D_{ic} is an ordinal variable, where τ stands for the choice category, and κ for the unknown threshold parameters on different levels of truly felt deprivation D^*_{ic} that are represented by the choice categories. Thus, D_{ic} is a positive monotonic transformation of the underlying latent variable of truly felt deprivation D^*_{ic} (see Maddala 1986; Greene 2010).

Second, due to the omission of explicit information on visibility of the objects of comparison in empirical survey data, we use relative deprivation in visible wealth DV_{ic} and in income DI_{ic} as proxies for D^v_{ic} and D^n_{ic} respectively, then

$$D^v_{ic} = \theta_1 DV_{ic} + \varepsilon_{1i}, \tag{2.5}$$

$$D^n_{ic} = \theta_2 DI_{ic} + \varepsilon_{2i}. \tag{2.6}$$

Based on the theoretical structure in equation 2.2, and the empirical proxies shown in equations 2.4 2.5 and 2.6, our empirical estimation strategy is as follows:

$$D_{ic} = \alpha + \beta_1 DV_{ic} + \beta_2 DI_{ic} + \sum_{k}^{K} \gamma_k x_{ki} + \sum_{l}^{L} \delta_l w_{lc} + \varepsilon_i \tag{2.7}$$

with the individual specific factors x_{ik} include sex, age, marriage, being a local in the own village or town. The group-specific characteristics w_{lc} distinguish between regional differences in the oblasts (provinces) and between rural and urban livelihoods.

In theory, we predict that only visible indicators have a significant role on the levels of perceived relative deprivation. Yet, given the implausibility in observational data to distinguish clearly between DV_{ic} and DI_{ic}, we hypothetically test for $\beta_1 > \beta_2 \geq 0$. If we are able not to reject this hypothesis, then we can confidentially conclude that visibility is an important aspect in measuring relative deprivation.

2.3 Measuring Visible Wealth and Deprivation

2.3.1 Data

We rely on the second wave of the 2011 *Life in Kyrgyzstan* (LIK) survey as it includes unique information on subjective deprivation in 120 towns and villages.[2] The socio-economic household survey is representative for the population in Kyrgyzstan using a stratified two-stage random sampling based on the 2009 Census (Brück et al. 2014).[3] For the analysis, we pool the heads of 2,809 households with valid information on perceived deprivation, income, consumption and asset items.[4] The non-response within the survey is generally low, with few missing values in perceived deprivation (less than 2%) and even fewer missing values in the economic indicators (less than 1%).

The key dependent variable and our benchmark for the comparison between visible and invisible deprivation is the question on self-assessed economic deprivation: *"How would*

[2]Attrition between the baseline in 2010 (with an original sample size of 3,000 households) and the follow-up wave in 2011 which is used for this paper is low (4.56%). The attrition households are mostly urban dwellers who are slightly, but not significantly, poorer than the households remaining in the second wave.

[3]The survey was conducted by the German Institute for Economic Research (DIW Berlin) in cooperation with local partners in Bishkek. The questionnaires in Kyrgyz, Russian and English can be accessed at the project website: *http://www.diw.de/kyrgyzstan.*

[4]If the head of the household was not available during the interview, we use information of the spouse or the most senior household member who responded.

you rate your household's current economic situation compared with other people in your town or village?" Respondents were given the option to place themselves on an 11-point Likert scale. The question was part of a set of life satisfaction questions with a common answer scale labeled as "completely dissatisfied" for category zero and "completely satisfied" for category ten. Given that the question aims to retrieve information on an individual's subjective *rating* within the town or village, rather than satisfaction, we are certain that the respondents gave information on their perceived relative deprivation. Individuals who perceived themselves as economically at the bottom of the town or village, chose category zero, while individuals who perceived themselves as at the top chose ten. The question has unique features that facilitate an in-depth analysis of the determinants of deprivation. First, the respondents reveal only their *perceived* level of deprivation relative to others. This kind of subjective information on deprivation comes presumably closer to the underlying feeling of deprivation than any calculated deprivation measure. Second, the question defines the comparison group, namely other people in the town or village, such that there is no need for making any artificial assumptions about with whom people compare themselves. In addition, the geographical unit of comparison is small enough to assure comparability between the in-group individuals, and hence strengthens the possibility to satisfy the visibility condition. To make the variable comparable to the calculated measures of deprivation (see below, section 2.3.3), the variable is inverted and normalized, with zero standing for "not deprived at all" and one for "completely deprived".[5]

For the calculation of deprivation in income, we use information on 22 different income sources of all household members, including monthly wages and salaries, social transfers, material aids, and income from household enterprises, from property and other income sources. The indicator of deprivation in visible wealth is constructed (see below, section 2.3.2) using information on a large range of asset and consumption items. These include the ownership of 40 different assets and the respective monetary values for 14 of them, as well as the monetary values for 21 non-food consumption items.

2.3.2 Measuring Visible Wealth

To compare the notion of visible wealth with income, a measure for visible wealth needs to be created. This implies several challenges. First, visible wealth is manifested in numerous assets and consumption goods. The visibility of an item is not a clear-cut

[5] This trivial transformation does not affect the underlying structure of the variable. It is undertaken just to provide easier interpretation of the coefficients of the actual deprivation measures in the analysis, where 0 signifies "not deprived at all" and 1 "fully deprived".

feature. Each of the items might meet the visibility criteria in some circumstances, while in others it does not. Furthermore, the items are usually measured in different scales, as for some, survey respondents are able to report the monetary values (continuous scale), while for others, only information on the ownership (binary or count scale) can be recalled.
From the pool of asset and consumption items in the LIK data, we identify 38 asset and consumption items (see table 2.1) that are usually observable by others within the village or town. To deal with the different scales, we aggregate first the selected variables within each of the dimensions, housing, transport, livestock, durables and consumption, before calculating relative deprivation and creating a composite indicator.[6]

For the aggregation of the variables within each dimension, we use weights derived from Principal Component Analysis (PCA), as they retain most of the variance of the original variables (Kolenikov and Angeles 2004; 2009). The concept of PCA is based on the idea that the N visible items v of household i are connected by underlying components C as follows:

$$\begin{aligned} v_{1i} &= \alpha_{11}C_{1i} + \alpha_{12}C_{2i} + \ldots + \alpha_{1N}C_{Ni}, \\ &\ldots \\ v_{Ni} &= \alpha_{N1}C_{1i} + \alpha_{N2}C_{2i} + \ldots + \alpha_{NN}C_{Ni}, \end{aligned}$$

with α_n as the eigenvectors. As only the visible goods v and not the components C are observed, these equations are inverted to find orthogonal linear combinations that retain a large part of the variance of the original variables. The first principal component consists of the sum of the items v multiplied with the PCA weights β:

$$C_{1i} = \beta_{11}v_{1i} + \beta_{12}v_{2i} + \ldots + \beta_{1N}v_{Ni}. \tag{2.8}$$

Accordingly, the magnitude of each weight depends on the extent of information that the item provides about the other items. For example, if expenditures in eating-out is highly correlated with other dimensions of visible consumption, such as the purchase of clothing, then the weight for eating-out consumption in the visible wealth index becomes large and positive. Within each dimension, we conduct principal component analysis on the selected variables. In each of the five dimensions, we find the weights from the

[6] Housing, means of transport and consumption items are assessed in monetary units (continuous variable), and durables in the quantities owned (count variable). For comparability, the livestock quantities are transferred into livestock equivalent units. We use the FAO (1982) equivalence scales (one horse 1 unit, one cow 1 unit, one sheep 0.15 units, one goat 0.15 units, one pig 0.15 units).

first component to be highly relevant for the outcome indicator of perceived deprivation.[7] Hence, the weights from the first component are taken when aggregating the selected variables.

2.3.3 Calculating Deprivation

There are a number of methodologies to calculate relative deprivation. The simplest and most widely used deprivation measure is the deviation from the group average, such as the arithmetic mean (e.g. Luttmer 2005; Bossert et al. 2013) or the geometric mean (e.g. Jasso and Wegener 1997). Other studies use rank-based measures to classify the relative position of individuals to the rest of the comparison group (Brown et al. 2008; Boyce et al. 2010). Powdthavee (2009) finds that the ordinal ranking in a group explains individual perception of economic standing better than the arithmetic mean. For this study, we choose the Yitzhaki (1979) measure of deprivation. The measure combines both mean- and rank-based measures, and is for each individual i calculated as follows:

$$D^y_{ic} = \sum_{j=i;c_i(y)}^{N-1} (1-p(y_i))(y_{j+1}-y_i) = \frac{1}{N}\sum_{j=i;c_i(y)}^{N-1} (y_{j+1}-y_i), \tag{2.9}$$

with wealth units $y_1, y_2, ..., y_n$ such that $y_1 \leq y_2 \leq ... \leq y_n$, $p(y) = prob(y_i \leq y_j)$ and comparison group c_i such that $c_i(y) = \left\{ j \in N \middle| y_j > y_i \right\}$. In other words, the level of deprivation of person i is calculated as the sum of differences between i's wealth and the wealth of all others above, divided by the total number of individuals N in the comparison group c. This implies that rank is important given that individuals only compare themselves against those in the comparison group who are better-off. The person with the highest wealth among all individuals in a comparison group has a value of $D^y_{ic} = 0$.

We calculate Yitzhaki measures of deprivation for income D^n_{ic} and for each of the different dimensions of visible wealth. All measures are normalized to a range from zero (not deprived at all) to one (completely deprived) to make them comparable. Based on the Yitzhaki measures in the different dimensions, a composite indicator of deprivation in visible wealth D^v_{ic} is calculated giving an equal weight for each dimension. This linear aggregation implies that a person who has less in one dimension can compensate in the other dimensions, as suggested by Permanyer (2014). The separate calculation of deprivation within each dimension of visible wealth implies two further assumptions.

[7] Higher principal components did not show significant correlations and, in some cases, not the expected negative direction of the relationship with perceived deprivation. Therefore we did not use higher components for the analysis.

First, comparisons are made within the dimensions, such that individuals compare their dwellings with the dwellings of others. Second, the deprivation of visible wealth of a household includes only dimensions with some variation within the town or village. For example, for households in the city without any livestock in the comparison group, the livestock dimension is not considered in the indicator for deprivation in visible wealth.

We choose the other sampled households in the town or village in the LIK survey as the comparison group in the calculation of deprivation. In the survey, 25 households were randomly selected in each primary sampling unit. Hence, the comparison group consists of 23 to 24 households in the town or village who were available in the 2011 interview and did not drop out after the first year of data collection. Although, we are aware of the fact that the sample is not representative at this level, the other sampled households within the town or village seem to be the most adequate comparison group. Other households in the town or village are within reach of the respondent. As such, they come closest both to Runciman's (1966) definition and to the specified comparison group in the question on perceived deprivation. As a check for the adequate comparison group, we calculated deprivation measures at the regional and the national level. We find the relationship between reported deprivation and calculated deprivation on the regional or national level is significantly less pronounced than on the town or village level which supports our choice of the comparison group.

2.4 Deprivation in Visible Wealth and Income Compared

2.4.1 Descriptives

One quarter of the respondents feel deprived and chose a category higher than the middle. This means that most individuals in our sample do not feel deprived. The average level of deprivation μ in visible wealth is 0.49. This is clearly the highest in comparison to deprivation in income ($\mu = 0.42$) and perceived deprivation ($\mu = 4$) (see table 2.2). At the same time, deprivation in visible wealth is much less volatile than deprivation in income and perceived deprivation. The coefficient of variation (cv)[8] is 0.41 for deprivation in visible wealth, while it is 0.69 for deprivation in income and 0.53 for perceived deprivation. Comparing the correlation ρ between perceived deprivation and the absolute deprivation measures, perceived deprivation is significantly higher correlated with deprivation in visible wealth ($\rho = 0.33$) than with deprivation income ($\rho = 0.24$).

[8] The cv is defined as the ratio of the standard deviation and the mean.

2.4.2 Main Results

We estimate deprivation in visible wealth and in income on perceived deprivation as specified in equation 2.7 using OLS[9] (see table 2.3). The standard errors are adjusted for clustering at the town or village level (Moulton 1990). In the joint estimation, we find that coefficients for both relative deprivation in income and visible wealth are relevant in explaining perceived deprivation. Furthermore, there is a notable and significant difference between both coefficients, where the visible wealth indicator captures the strongest effect. To be more specific, the visible wealth indicator explains perceived deprivation almost threefold more than income (column 1, table 2.3). In order to provide better comparable clarification of the effect between income and visible wealth, we run separate regressions (columns 2-3, table 2.3). Both coefficients retain a highly significant and positive effect. A unit increase in actual relative deprivation in visible wealth increases perceived deprivation by more than three Likert-scale points. A unit increment in deprivation in income on the other hand, only increases it by 1.5 Likert-scale points.[10] Although both income and visible wealth play a potent role in explaining perceived deprivation, the visibility of the object of comparison adds more certainty in enabling the relative economic assessment of individuals. On another note, it is important to point out that urban dwellers perceive themselves to be more deprived than rural dwellers. Moreover, the more educated people are, the less deprived they perceive themselves to be. This is evident from the negative significant coefficient for all the specifications and may reflect the expected perception and the potential of the highly educated population to be economically better off.

2.4.3 Robustness Checks

We perform a number of checks in order to test for the robustness of the results. These robustness and sensitivity tests aim to mitigate any concerns arising from structural biases of both the main dependent and independent variables. First, we replace the Yitzhaki measure with rank and mean measures. Second, we rerun the regressions for each sub-index of the visible wealth index. This helps us identify the strength of the composite index. Third, we use equal weights instead of principal component weights in addition to separately calculated PCA weights for rural and urban dwellers in the visible wealth index. Fourth, we perform the analysis excluding the middle category of the dependent variable given the high selection share of the total respondents (25%) for this category, which stands neither for relative deprivation nor for relative satisfaction. In short, we

[9]Given that the dependent variable is an 11-point Likert scale, it can be treated as a continuous variable. Hence, a simple OLS estimation suffices. Results do no change when using Ordered Logit estimations.

[10]A unit increase represents a shift from no actual deprivation to complete deprivation within a town or village.

do not find any noteworthy deviation from the main result and can confidently conclude that the visible wealth index plays a more important role than income in determining perceived levels of relative deprivation.[11]

The first test relates to our choice of the Yitzhaki deprivation measure to generate values on objective (actual) relative deprivation. Another branch of the literature uses other approaches to calculate relative economic differences (see section 2.3.3). Two notable examples are mean-based (e.g. Luttmer 2005, Jasso and Wegener 1997) and rank-based measures (e.g. Boyce et al. 2010). We calculate the rank, as well as the arithmetic and geometric means of income and visible wealth within each village and town in our sample (see table 2.4).[12] In line with the literature, the mean-based measures do not significantly explain perceived deprivation, yet the absolute levels of both income and visible wealth do. More importantly, we find that the rank-based measure of visible wealth has a negative and significant effect on perceived deprivation, while the coefficient of the income rank is insignificant. This further stresses the importance of visible wealth in contrast to income for explaining feelings of deprivation.

Second, we rerun the estimation using the sub-indices of visible wealth instead of the aggregated index in order to check if all sub-categories are imperative to include in the construction of visible wealth indicator. Each sub-index is significant when entered as the only objective deprivation predictor after controls (see table 2.5). This result shows that all five dimensions have the expected positive and significant sign, despite some variation in the size of the effect. All sub-categories of the visible wealth index are relevant for explaining perceived deprivation.

Third, we address a common concern related to principal components and check whether the PCA weights fundamentally drive the results by running estimations using equal weights for all items within a dimension (columns 1-3, table 2.6). The results basically remain unchanged. Both coefficients of deprivation in visible wealth and in income are significantly positive and different from each other. Furthermore, we calculate principal component weights separately for rural and urban dwellers in order to account for any structural differences in the sample (columns 4-6, table 2.6). This exercise takes account of the fact that e.g. livestock is considered to be an indicator of wealth in the countryside,

[11] We further run regressions including only the households in the 99 percentile of income, assets and consumption. We find the results robust to outliers.

[12] As the mean- and rank-based measures only marginally include the own level of economic wealth, an individual's absolute income (log) and the absolute level of PCA score in visible wealth are included as additional controls in the estimations.

while it is a sign of being poor in the cities. The results support again our findings from the baseline regressions.

A large portion of the respondents (25%) chose the middle category (five) of perceived deprivation from the 11-Likert scale. Given that this choice does neither reflect relative deprivation nor relative satisfaction, individuals may use it as an undetermined answer choice. Moreover, some individuals who are completely deprived or not deprived at all may also place themselves in category five. These people might not want to reveal their true perceived relative status. In order to understand the underlying factors behind the high rate of selection of the middle category and its effect on the primary results, we further conduct different sensitivity checks. First, we restrict our sample by excluding the individuals who chose the middle category of subjective deprivation (columns 1-2, table 2.7). Second, we perform a lower and upper bounding replacement of the category five. In the lower (upper) bound scenario, we assume that all respondents of the middle category do not want to reveal themselves as better-off or worse-off. Accordingly, we allocated the middle category respondents in the lower bound scenario to the non-deprived ($D_{ic} = 4$) and in the upper bound scenario to the deprived ($D_{ic} = 6$) and run separate regressions (columns 3-6, table 2.7). The results in all specifications remain similar to the baseline.

2.5 Conclusion

Motivated by the theory on deprivation, we expect visible wealth to be more important than income for somebody's feeling of deprivation. To test this hypothesis, we use household data from Kyrgyzstan with information on perceived deprivation. After identifying a range of visible items and calculating an indicator of deprivation in visible wealth, we evaluate the relationship of deprivation in visible wealth and that of income with perceived subjective deprivation. We find that deprivation in visible wealth plays an important role in explaining perceived deprivation. The effect of deprivation in visible wealth on perceived deprivation is significantly stronger than that of the standard measure of deprivation in income. Our finding is robust under various sensitivity checks and for a number of controls.

The results from the study shed light on the role of visible wealth for relative deprivation. There is still room to develop and improve the measure of visible wealth by better classifying visible wealth items, or calculating scores and weights for the composite measure. Nevertheless, the paper should be understood as the starting point for further research on deprivation and its impacts on human well-being and behavior to not just rely on income as a measure for relative deprivation, but also focus on visible wealth, given its indispensable role on an individual's perception of deprivation. First important steps would be to collect more detailed information on the visibility and values of goods and assets, on the relevant comparison group and perceived deprivation in different dimension of well-being. Last but not least, researchers should exploit the whole range of information on goods and assets available, when calculating deprivation measures based on socio-economic survey data.

Tables

Table 2.1: Visible asset items and goods

variable	mean	sd	N
main dwelling	977,787.78	788,945.00	2,859
another house or apartment	28,056.26	196,858.63	2,859
motorcycle, scooter	156.26	2,588.94	2,859
car, minibus	62,309.70	289,665.92	2,859
tractor, truck	12,166.37	102,035.02	2,859
cow, bull	1.15	1.86	2,859
sheep, goat	4.47	12.91	2,859
horse	0.19	0.89	2,859
pig	0.04	0.52	2,859
chicken	4.59	9.09	2,859
fridge	0.80	0.41	2,859
electric stove	0.71	0.46	2,859
microwave	0.28	0.45	2,859
air conditioner	0.04	0.21	2,859
sewing machine	0.54	0.50	2,859
washing machine (automatic)	0.19	0.40	2,859
vacuum cleaner	0.52	0.50	2,859
sofa	0.96	0.58	2,859
wardrobe	1.17	0.68	2,859
bed	1.20	1.11	2,859
kitchen furniture	0.29	0.46	2,859
radios	0.14	0.35	2,859
music systems	0.12	0.33	2,859
television	1.16	0.49	2,859
video player	0.81	0.46	2,859
video camera	0.03	0.18	2,859
photo camera	0.05	0.23	2,859
photo camera (digital)	0.07	0.27	2,859
computer	0.10	0.32	2,859
satellite dish	0.18	0.39	2,859
mobile phone	1.54	0.92	2,859
entertainment, recreation, eating out	309.50	627.24	2,861
internet, cable tv, communication	331.25	355.87	2,861
celebration, funerals, rituals	613.85	1,353.04	2,861
education expenses	191.62	646.32	2,861
clothing and shoes	876.88	790.99	2,861
furniture and other interiors	129.12	590.41	2,861
other durable goods	120.09	1,203.47	2,861

Data Source: Life in Kyrgyzstan Survey 2011.

Table 2.2: Summary statistics

variable	mean	sd	min	max	N
deprivation measures					
perceived deprivation	0.41	0.22	0	1	2,824
deprivation in income	0.43	0.30	0	1	2,855
deprivation in visible wealth (pca wt.)	0.49	0.20	0	1	2,863
individual-specific characteristics					
sex (dummy, 1=male, 0=female)	0.71	0.45	0	1	2,863
age	51.23	14.15	16	99	2,863
married (dummy)	0.71	0.46	0	1	2,863
Kyrgyz (dummy)	0.67	0.47	0	1	2,863
born in this town/village (dummy)	0.76	0.43	0	1	2,858
education level	4.77	1.44	1	7	2,858
community-specific characteristics					
urban (dummy)	0.41	0.49	0	1	2,863
Issyk-Kul oblast (dummy)	0.09	0.29	0	1	2,863
Djalal-Abad oblast (dummy)	0.16	0.37	0	1	2,863
Naryn oblast (dummy)	0.04	0.21	0	1	2,863
Batken oblast (dummy)	0.08	0.27	0	1	2,863
Osh oblast (dummy)	0.17	0.37	0	1	2,863
Talas oblast (dummy)	0.04	0.20	0	1	2,863
Chui oblast (dummy)	0.17	0.37	0	1	2,863
Bishkek oblast (dummy)	0.20	0.40	0	1	2,863
Osh oblast (dummy)	0.04	0.21	0	1	2,863

Data Source: Life in Kyrgyzstan Survey 2011.

Table 2.3: Determinants of perceived deprivation comparing deprivation in income and visible wealth

	(1)	(2)	(3)
dependent variable	perceived deprivation[1]	perceived deprivation[1]	perceived deprivation[1]
independent variables			
deprivation in income[2]	0.863***	1.547***	
	(0.204)	(0.194)	
deprivation in visible wealth (pca wt.)[23]	2.579***		3.087***
	(0.357)		(0.333)
urban (dummy)	0.638**	0.737**	0.602*
	(0.317)	(0.323)	(0.320)
sex (dummy)	0.038	-0.012	0.041
	(0.148)	(0.151)	(0.149)
age	-0.001	-0.003	-0.002
	(0.004)	(0.004)	(0.004)
married (dummy)	-0.114	-0.238*	-0.207
	(0.145)	(0.143)	(0.151)
Kyrgyz (dummy)	-0.261	-0.278	-0.281
	(0.220)	(0.226)	(0.219)
born in this town/village (dummy)	0.100	0.039	0.116
	(0.196)	(0.197)	(0.195)
education level	-0.122***	-0.183***	-0.130***
	(0.043)	(0.042)	(0.043)
constant	3.304***	4.942***	3.574***
	(0.841)	(0.817)	(0.831)
oblast dummies	YES	YES	YES
adj. R2	0.166	0.124	0.156
observations	2809	2809	2809

Notes: All estimations are based on OLS regressions. The standard errors are clustered on the level of a village or town and reported in brackets; *** p<0.01; ** p<0.05; *p<0.1 significance level.

1 Perceived deprivation reflects the subjective rating of the household head with respect to the economic situation of the household in comparison to others within the town or village and is based on a 11-Likert scale question. The original variable is inverted and standardized, ranging from zero (not at all deprived) to one (completely deprived).

2 Relative deprivation is calculated according to the Yitzhaki deprivation formula using all other surveyed households within a village or town as the comparison group. The variable ranges from zero (not at all deprived) to one (completely deprived).

3 The visible wealth index includes 38 different variables of visible asset and consumption items. Within the five dimensions, housing, transport, livestock, durables and consumption, visible items were aggregated using PCA weights. Each dimension entered with an equal weight the visibility index which is used as the basis for calculating deprivation in visible wealth within the town or village.

Table 2.4: Determinants of perceived deprivation comparing rank- and mean-based deprivation measures in income and visible wealth

	(1)	(2)	(3)	(4)	(5)	(6)
dependent variable	perceived deprivation[1]	perceived deprivation[1]	perceived deprivation[1]	perceived deprivation[1]	perceived deprivation[1]	perceived deprivation[1]
independent variables						
income (log)	-0.309		-0.656***		-0.646***	
	(0.260)		(0.088)		(0.085)	
rank income[2]	-0.768					
	(0.513)					
visible wealth index (pca score)[4]		-0.277**		-0.464***		-0.468***
		(0.106)		(0.062)		(0.072)
rank visible wealth[24]		-0.851***				
		(0.252)				
mean income (log)[3]			0.419			
			(0.379)			
mean visible wealth (pca score)[34]				0.105		
				(0.227)		
geometric mean income (log)[3]					0.342	
					(0.383)	
geometric mean visible wealth (pca score)[34]						0.596
						(0.399)
urban (dummy)	0.714**	0.640*	0.716**	0.627*	0.717**	0.680**
	(0.335)	(0.327)	(0.332)	(0.327)	(0.336)	(0.339)
sex (dummy)	0.001	0.018	-0.007	0.010	-0.008	0.000
	(0.150)	(0.152)	(0.150)	(0.152)	(0.150)	(0.154)
age	-0.003	-0.002	-0.003	-0.003	-0.003	-0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
married (dummy)	-0.260*	-0.289*	-0.256*	-0.323**	-0.260*	-0.326**
	(0.146)	(0.148)	(0.145)	(0.148)	(0.145)	(0.150)
Kyrgyz (dummy)	-0.286	-0.298	-0.288	-0.299	-0.285	-0.290
	(0.220)	(0.219)	(0.221)	(0.218)	(0.219)	(0.221)
born in this town/village (dummy)	0.058	0.093	0.056	0.098	0.058	0.121
	(0.195)	(0.188)	(0.197)	(0.188)	(0.197)	(0.197)
education level	-0.177***	-0.150***	-0.178***	-0.165***	-0.178***	-0.155***
	(0.043)	(0.045)	(0.043)	(0.044)	(0.043)	(0.044)
constant	8.821***	5.819***	7.694**	5.547***	8.410**	5.095***
	(2.220)	(0.843)	(3.620)	(0.836)	(3.530)	(0.862)
oblast dummies	YES	YES	YES	YES	YES	YES
adj. R2	0.116	0.122	0.118	0.124	0.117	0.128
observations	2809	2809	2809	2809	2809	2761

Notes: All estimations are based on OLS regressions. The standard errors are clustered on the level of a village or town and reported in brackets; *** p<0.01; ** p<0.05; *p<0.1 significance level.

[1] Perceived deprivation reflects the subjective rating of the household head with respect to the economic situation of the household in comparison to others within the town or village and is based on a 11-Likert scale question. The original variable is inverted and standardized, ranging from zero (not at all deprived) to one (completely deprived).

[2] The rank is calculated based on the household's position within the community using all other surveyed households in the town or village. The variable is standardized ranging from zero (lowest rank) to one (highest rank).

[3] The mean is based on information of all other surveyed households within the town or village of the responding household.

[3] The visible wealth index includes 38 different variables of visible asset and consumption items. Within the five dimensions, housing, transport, livestock, durables and consumption, visible items were aggregated using PCA weights. Each dimension entered with an equal weight the visibility index which is used as the basis for calculating deprivation in visible wealth within the town or village.

Table 2.5: Determinants of perceived deprivation comparing deprivation in income and different dimensions of visible wealth

	(1)	(2)	(3)	(4)	(5)
dependent variable	perceived deprivation[1]	perceived deprivation[1]	perceived deprivation[1]	perceived deprivation[1]	perceived deprivation[1]
independent variables					
deprivation in income[2]	1.411***	1.222***	1.356***	1.308***	1.083***
	(0.191)	(0.198)	(0.223)	(0.209)	(0.208)
deprivation in housing values (pca wt.)[23]	0.842***				
	(0.212)				
deprivation in transport values (pca wt.)[23]		1.173***			
		(0.171)			
deprivation in livestock (pca wt.)[23]			0.501**		
			(0.218)		
deprivation in durables (pca wt.)[23]				0.907***	
				(0.204)	
deprivation in consumption (pca wt.)[23]					1.125***
					(0.258)
urban (dummy)	0.742**	0.722**	0.503	0.704**	0.747**
	(0.320)	(0.321)	(0.317)	(0.321)	(0.315)
sex (dummy)	-0.037	0.109	-0.041	-0.012	-0.015
	(0.152)	(0.152)	(0.178)	(0.150)	(0.146)
age	-0.001	-0.004	-0.003	-0.002	-0.006
	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)
married (dummy)	-0.178	-0.198	-0.136	-0.224	-0.189
	(0.145)	(0.141)	(0.158)	(0.145)	(0.139)
Kyrgyz (dummy)	-0.281	-0.290	-0.375	-0.314	-0.213
	(0.226)	(0.219)	(0.251)	(0.227)	(0.227)
born in this town/village (dummy)	0.080	-0.004	-0.116	0.060	0.059
	(0.197)	(0.197)	(0.245)	(0.197)	(0.199)
education level	-0.164***	-0.149***	-0.144***	-0.150***	-0.156***
	(0.044)	(0.043)	(0.039)	(0.042)	(0.043)
constant	4.341***	3.989***	5.546***	4.451***	4.576***
	(0.837)	(0.811)	(0.813)	(0.811)	(0.833)
oblast dummies	YES	YES	YES	YES	YES
adj. R2	0.135	0.158	0.142	0.136	0.141
observations	2807	2738	2281	2809	2809

Notes: All estimations are OLS estimations. The standard errors are clustered on the level of a village or town and reported in brackets; *** p<0.01; ** p<0.05; *p<0.1 significance level.

[1] Perceived deprivation reflects the subjective rating of the household head with respect to the economic situation in comparison to others within the town or village and is based on a 11-Likert scale question. The variable is inverted and standardized, ranging from zero (not at all deprived) to one (completely deprived).

[2] Relative deprivation is calculated according to the Yitzhaki deprivation formula using all other surveyed households within a village or town as the comparison group. The variable ranges from zero (not at all deprived) to one (completely deprived).

[3] The individually reported items were aggregated within the dimension by using PCA weights.

Table 2.6: Determinants of perceived deprivation comparing deprivation in income and deprivation in visible wealth using different weights for the visibility index

	(1)	(2)	(3)	(4)	(5)
dependent variable	perceived deprivation[1]	perceived deprivation[1]	perceived deprivation[1]	perceived deprivation[1]	perceived deprivation[1]
independent variables					
deprivation in income[2]	0.856***	1.547***		0.869***	
	(0.202)	(0.194)		(0.197)	
deprivation in visible wealth (equ. wt.)[234]	2.561***		3.069***		
	(0.350)		(0.329)		
deprivation in visible wealth (pca wt., by location)[235]				2.577***	3.089***
				(0.382)	(0.362)
urban (dummy)	0.488	0.737**	0.466	0.594*	0.549*
	(0.308)	(0.323)	(0.308)	(0.313)	(0.316)
sex (dummy)	-0.005	-0.012	-0.005	0.030	0.032
	(0.178)	(0.151)	(0.180)	(0.147)	(0.149)
age	-0.000	-0.003	-0.002	-0.001	-0.002
	(0.005)	(0.004)	(0.005)	(0.004)	(0.004)
married (dummy)	-0.033	-0.238*	-0.101	-0.108	-0.201
	(0.156)	(0.143)	(0.162)	(0.145)	(0.150)
Kyrgyz (dummy)	-0.373	-0.278	-0.383	-0.263	-0.283
	(0.243)	(0.226)	(0.240)	(0.220)	(0.218)
born in this town/village (dummy)	-0.064	0.039	-0.064	0.096	0.112
	(0.238)	(0.197)	(0.238)	(0.196)	(0.196)
education level	-0.086**	-0.183***	-0.095**	-0.124***	-0.133***
	(0.041)	(0.042)	(0.041)	(0.043)	(0.044)
constant	4.155***	4.942***	4.390***	3.400***	3.692***
	(0.841)	(0.817)	(0.843)	(0.856)	(0.849)
oblast dummies	YES	YES	YES	YES	YES
adj. R2	0.166	0.124	0.156	0.166	0.155
observations	2809	2809	2809	2809	2809

Notes: All estimations are based on OLS regressions. In columns (1) and (2), estimations are based on a sub-sample without middle category respondents. In columns (3) and (4), middle category respondents were considered as deprived with a deprivation value of 0.6, in columns (5) and (6), the middle category respondents were considered as not deprived with a deprivation value of 0.4. The standard errors are clustered on the level of a village or town and reported in brackets; *** p<0.01; ** p<0.05; *p<0.1 significance level.

1 Perceived deprivation reflects the subjective rating of the household head with respect to the economic situation of the household in comparison to others within the town or village and is based on a 11-Likert scale question. The original variable is inverted and standardized, ranging from zero (not at all deprived) to one (completely deprived).

2 Relative deprivation is calculated according to the Yitzhaki deprivation formula using all other surveyed households within a village or town as the comparison group. The variable ranges from zero (not at all deprived) to one (completely deprived).

3 The visible wealth index includes 38 different variables of visible asset and consumption items within five dimensions. Each dimension entered the index with an equal weight. The visible wealth index is taken as the basis for calculating deprivation in visible wealth of a household within the town or village.

4 In each dimension, the variables of visible asset and consumption items are aggregated using equal weights.

5 In each dimension, the variables of visible asset and consumption items are aggregated using PCA weights that were separately calculated for rural and urban neighborhoods.

Table 2.7: Determinants of perceived deprivation comparing deprivation in income and visible wealth: sensitivity analysis on middle category respondents

	(1)	(2)	(3)	(4)	(5)	(6)
dependent variable	perceived deprivation[1]	perceived deprivation[1]	perceived deprivation[1]	perceived deprivation[1]	perceived deprivation[1]	perceived deprivation[1]
independent variables						
deprivation in income[2]	1.985***		1.568***		1.526***	
	(0.233)		(0.207)		(0.184)	
deprivation in visible wealth (pca wt.)[2,3]		3.915***		3.213***		2.961***
		(0.380)		(0.354)		(0.320)
urban (dummy)	0.973**	0.860**	0.722**	0.583*	0.752**	0.621**
	(0.409)	(0.411)	(0.352)	(0.346)	(0.302)	(0.302)
sex (dummy)	-0.062	0.055	-0.001	0.054	-0.023	0.028
	(0.191)	(0.186)	(0.159)	(0.156)	(0.147)	(0.147)
age	-0.003	-0.001	-0.004	-0.003	-0.003	-0.002
	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)
married (dummy)	-0.273	-0.295	-0.221	-0.182	-0.255*	-0.232
	(0.194)	(0.196)	(0.145)	(0.153)	(0.146)	(0.153)
Kyrgyz (dummy)	-0.275	-0.294	-0.310	-0.312	-0.246	-0.250
	(0.263)	(0.252)	(0.250)	(0.242)	(0.205)	(0.199)
born in this town/village (dummy)	-0.003	0.093	0.073	0.153	0.006	0.080
	(0.237)	(0.235)	(0.210)	(0.208)	(0.187)	(0.186)
education level	-0.217***	-0.159***	-0.188***	-0.132***	-0.178***	-0.128***
	(0.050)	(0.051)	(0.046)	(0.046)	(0.040)	(0.041)
constant	4.767***	3.102***	5.155***	3.697***	4.729***	3.452***
	(0.983)	(0.976)	(0.851)	(0.872)	(0.796)	(0.806)
oblast dummies	YES	YES	YES	YES	YES	YES
adj. R2	0.160	0.202	0.113	0.145	0.127	0.156
observations	2091	2091	2809	2809	2809	2809

Notes: All estimations are based on OLS regressions. In columns (1) and (2), estimations are based on a sub-sample without middle category respondents. In columns (3) and (4), middle category respondents were considered as deprived with a deprivation value of 0.6, in columns (5) and (6), the middle category respondents were considered as not deprived with a deprivation value of 0.4. The standard errors are clustered on the level of a village or town and reported in brackets; *** $p<0.01$; ** $p<0.05$; *$p<0.1$ significance level.

1 Perceived deprivation reflects the subjective rating of the household head with respect to the economic situation of the household in comparison to others within the town or village and is based on a 11-Likert scale question. The original variable is inverted and standardized, ranging from zero (not at all deprived) to one (completely deprived).

2 Relative deprivation is calculated according to the Yitzhaki deprivation formula using all other surveyed households within a village or town as the comparison group. The variable ranges from zero (not at all deprived) to one (completely deprived).

3 The visible wealth index includes 38 different variables of visible asset and consumption items. Within the five dimensions, housing, transport, livestock, durables and consumption, visible items were aggregated using PCA weights. Each dimension entered with an equal weight the visibility index which is used as the basis for calculating deprivation in visible wealth within the town or village.

Chapter 3

Attitudes toward Risk: An Evaluation of Elicitation Methods in the Field

Veronika Bertram-Hümmer
Karlijn Morsink

3.1 Introduction

Risk preferences are important to understand individual behavior and decisions under risk. The goal of this study is to enhance the measurement of risk preferences of low-educated inhabitants in developing countries who are highly vulnerable to shocks and exposed to risks. We utilize individual choices from an artefactual field experiment in rural Ethiopia to investigate two widely used methods to elicit information on risk preferences: the simple *Ordered Lottery Selection* of *Binswanger* (1980; 1981) (OLS-BW hereafter) and a more comprehensive *Multiple Price List* design (MPL hereafter) with repeated choices between a risky option and a safe amount.

Having explicit and reliable information on risk preferences is important when analyzing decisions and behavior under risk. Whether or not someone loves taking risks, and embraces the opportunities involved, will directly affect his or her behavior and can have indirect implications on welfare and poverty. There is evidence that risk preferences influence financial and employment decisions: the income of risk averse individuals is lower and can be characterized by slower growth (McInish et al. 1993; Shaw 1996), the risk averse are less likely self-employed (Hardeweg et al. 2013), and they are more hesitant to migrate abroad, which would otherwise lead to better labor and income opportunities (Jaeger et al. 2010). There is further evidence that formal and informal insurance decisions are influenced by attitudes toward risk (for example Giné et al. 2007; Giesbert et al. 2011; Attanasio et al. 2012). Risk preferences are also found to affect investment and production decisions: individuals who are generally risk averse adopt new technologies later on in the technology diffusion process, while individuals characterized by risk aversion to low probability risks adopt these earlier (Liu 2013). Any analysis of these kinds of decisions should benefit from adequate measures of risk preferences.

Consequently, risk preferences should be taken into account when designing and developing policy instruments that support individuals in their risk management. For example, when designing an insurance scheme, individual risk preferences might be helpful to derive reliable predictions on whether the target group would buy the insurance and how much they would be willing to spend (Carter et al. 2015).[1] When the risk preference

[1]To give a more detailed example, we assume the introduction of an insurance for farmers against harvest losses. Under the assumption that a farmer faces good weather with a likelihood of 70% and an expected return of 60 units, and bad weather with a likelihood of 30% and an expected return of 30 units, the farmer would have an expected utility without insurance of EU=0.7*60+0.3*30=51, implying a significant amount of risk. Under the assumption that the farmer insures the harvest for a premium of 15 units with indemnity payments of 30 units in case of bad weather, his or her expected utility would be EU_{ins}=0.7*60+0.3*(30+30)-15=60-15=45 and involve no risk. Under this scenario, risk averse farmers will

measures used in the underlying analysis of the policy design are imprecise or biased, the inferences drawn on people's behavior might be incorrect (Holt and Laury 2002) and the program might not meet the needs of the target group and fail its intended goals.

To account for attitudes toward risk when analyzing behavior, risk preferences need to be adequately elicited. This is not a trivial task: Unlike easily quantifiable information, for example a person's income, measuring risk preferences is not straightforward. In the classical economic theory, risk preferences have traditionally been considered to be an underlying trait (for example von Neumann and Morgenstern 1953). More recent literature finds risk preferences to be highly context-specific, such as the area of risk-taking (for example Weber et al. 2002; Anderson and Mellor 2009) or the risk environment (Harrison and Rutström 2008; Harbaugh et al. 2010). There is no single established method of how to adequately elicit and measure risk preferences (Charness et al. 2013). Elicitation methods differ in terms of the way they measure risk preferences and in terms of the level of noise they generate. Sophisticated methods allow for a fine characterization of risk attitudes and the estimation of different notions of risk attitudes, for example taking account of subjective probability weighting, but they might come at the cost of a lower comprehension by the individuals, causing higher levels of noise in the elicited measures (Dave et al. 2010). As a consequence, different elicitation methods often result in measures that reflect different risk preferences (e.g. Chuang and Schechter 2015). In a review, Charness et al. (2013) point out the prevalent knowledge gap on the effectiveness of the different risk preference elicitation methods. The few studies that empirically evaluate elicitation methods are based on experiments in developed countries (see Dave et al. 2010 and Crosetto and Filippin 2013). These studies focus on elicitation methods designed for standard samples in developed countries, that might be too complex and hence unsuitable for samples in developing countries characterized by low education levels and a large exposure to risk.

We investigate two widely used risk preference elicitation experiments from a sample of Ethiopian farmers. The OLS-BW experiment is based on a simple choice between six lotteries and is, as such, a fast and simple way to elicit risk preferences. It is generally the default way to elicit risk preferences in developing countries (see for example Yesuf and Bluffstone 2009; Barr et al. 2012; Attanasio et al. 2012). The MPL experiment is more comprehensive, requiring numerous binary choices from each participant and is increasingly used in developing countries (e.g. Henrich and McElreath 2002; Callen et al. 2014; Vieider et al. 2015). We follow methodologically the study of Dave et al. (2010),

theoretically buy the insurance, while risk neutral or risk loving farmers will not (example adopted from Carter et al. 2015).

who evaluate two common risk elicitation methods designed for standard samples in developed countries and analyze both non-parametric and parametric measures of risk preferences and their ability to reflect risk preferences in different risk environments. This comprises an investigation of estimated stochastic noise and non-parametric noise as measured by violations of stochastic dominance. Our estimations of risk preference measures and noise are based on Expected Utility Theory (EUT) and Rank-dependent Utility (RDU) theory using maximum likelihood techniques.

This paper contributes methodologically to the literature on risk preferences. To the best of our knowledge, it is the first study that empirically investigates risk preference elicitation methods tailored for low-educated samples in developing countries. Our study exploits risk choices from a sample of farmers in rural Ethiopia in both the OLS-BW and the MPL experiment. In contrast to studies that evaluate merely non-parametric measures of different methods (e.g. Loomes and Pogrebna 2014), our analysis builds on the parametric estimation of structural decision-making models. This allows us to disentangle risk preferences from stochastic noise and to investigate the interactions between risk preferences and wealth, as well as the level of risk exposure. In comparison to studies based on hypothetical tasks (e.g. De Brauw and Eozenou 2014), our analysis relies on incentivized tasks that are found to generate more reliable risk preference measures (Kachelmeier and Shehata 1992). Finally, the experimental approach allows us to isolate risk preferences within a laboratory environment, which is not feasible with non-experimental methods where the elicited risk preferences might only be valid within a certain context (Harrison and Rutström 2008).

We find that both measures reveal similar levels of risk preferences with moderate levels of stochastic noise when parametrically estimated. The estimated risk preference measures characterize our sample of Ethiopian farmers as risk loving. We further find decreasing relative risk aversion (*drra* preferences), which implies risk lovingness is increasing relative to wealth. In contrast to the design of the OLS-BW experiment, the MPL allows for the detection of moderate levels of inconsistent choices from our sample and to estimate more complex RDU models testing for subjective probability weighting. We find that subjects in our sample overweight low probability risks and underweight high probability risks (as indicated by an inverse s-shaped probability weighting function). Due to the cap at risk neutrality in the OLS-BW design, we find the non-parametric OLS-BW risk preference measure becomes heavily distorted toward risk aversion. When investigating the drivers of stochastic noise, we find noise in the MPL choices is significantly higher when the game is played in the afternoon (rather than the morning) sessions and that the enumerators significantly influence levels of noise in both experiments. Our findings suggest that the simple OLS-BW experiment is sufficient when parametrically estimating risk preferences, while the MPL experiment is preferable when relying on non-parametric methods to analyze risk preferences, when investigating inconsistent choices, and when estimating more complex economic models to analyze risk preferences in different risk environments.

The rest of the paper is structured as follows. Section 3.2 describes the design and the implementation of the OLS-BW and the MPL experiments, including a discussion of the implications of different design features on the risk preference measures to be elicited. Our methodological approach on the calculation of the non-parametric measures and the estimation of parametric risk preferences is explained in section 3.3. Section 3.4 comprises the empirical investigation of risk preference measures and noise based on the choices of our Ethiopian sample in the two experiments. Section 3.5 concludes.

3.2 Experimental Design

We evaluate risk preference elicitation methods based on two artefactual experiments in the Tigray region of Ethiopia. We played the OLS-BW and the MPL experiment with a sample of 875 individuals in 28 sessions between February 27 and March 28, 2014.[2] The experiments were incentivized and the final payoff was determined by a random lottery incentive mechanism (see section 3.2.2). Half of the sessions were played in the morning, and the other half in the afternoon. Before each session, we randomly assigned different orders between the OLS-BW and the MPL experiment, as well as between the price lists in the MPL experiment.[3] After the experiments, the participants responded to a small socio-economic survey that included questions on their willingness to take risks, which are used as complementary information for the evaluation of the experimental risk preference measures.[4]

[2]Each participant further participated in a risk sharing experiment that was held before the risk preference games.

[3]We are not able to control for order effects in the estimations, as the paperwork on the assignment of the order was lost during field work. Under the assumption that the order affects choices in the OLS-BW and MPL in the same way, the randomization procedure generates risk preferences and noise which are not driven by any order effects (as half of the sample played first the OLS-BW, and the other half played first the MPL). In case one of the two experiments is more sensitive to order, we cannot exclude the possibility that the order influences the estimated levels of risk preferences and noise.

[4]The survey consists of a general willingness-to-take-risk question, as well as questions about the willingness-to-take-risk when taking a loan, when buying seedlings, when buying fertilizer, when negotiating about crop prices and when negotiating about shared cropping.

3.2.1 Elicitation Design

Design Features

OLS-BW Design We conducted a simple ordered lottery selection experiment as introduced by Binswanger (1980; 1981). The OLS-BW design is a fast and easy way to elicit risk preferences and commonly used with samples in developing countries. Subjects make one decision between six lotteries that differ in terms of average payoffs and the variance around the payoffs (see table 3.1). In all lotteries, the participants have the chance of winning with a fixed probability of 50%. The first lottery does not involve any risk and can be considered as a safe amount, while risk is increasing over the remaining lotteries. The Expected Values (EV) in the lotteries are between 40 and 80 Ethiopian Birr (ETB).[5] The OLS-BW experiment took about 15 minutes within a session.

MPL Design The MPL format is a more comprehensive approach to elicit risk preferences as subjects make numerous binary choices between a safe amount and a lottery (see for example Bruhin et al. 2010; Callen et al. 2014). In our MPL design, respondents make choices within six price lists, reflecting different risk environments (see table 3.2).[6] The lotteries include winning and losing lotteries with three different risk probabilities (25%, 50% and 75%). Each price list incorporates 17 decision rows in which the lottery stays the same, while the safe amount increases over the rows (see figure 3.1). The EVs in the lotteries of the different price lists are between -120 and 120 ETB. Our participants answered 102 binary questions in the six price lists, which took about 45 minutes.

Implications of the Design

The elicitation design of an experiment plays a fundamental role on how individuals make their choices. It determines *a priori* the risk preference measures to be elicited as well as how precisely and which dimensions of risk preferences can be measured (Charness et al. 2013). Furthermore, the design determines the complexity of the experiment and affects the likelihood that individuals will not be able to understand the tasks. If an experiment is highly complex, subjects are more likely to have difficulties in understanding, to get easily tired and to be unable to make the choices that reveal their true risk preferences. This concern is particularly relevant when - as in our case - risk preferences are elicited from

[5] The exchange rate was 1 US$ = 8.28 ETB at the time of the experiments (in March 2014).

[6] The experiment also included a price list with a mixed lottery in which respondents could win or lose money, both with a probability of 50%. We did not use the mixed lottery in this paper as our sample had extreme difficulties in understanding and responding: approximately 67% (or 580) of the respondents never switched and chose the lottery option in all rows and 11 subjects were not able to make any choice at all. Our experience is similar to that of Yesuf and Bluffstone (2009, p. 1026) in Ethiopia.

a low-skilled sample (Dave et al. 2010). In the following, we discuss the implications of the design features in the OLS-BW and the MPL that need to be taken into account when analyzing and interpreting the elicited risk preference measures.

Implications on Risk Preferences Measures We identify three features in the design of the experiments that directly affect the measures to be elicited. First, the range of the outcomes in the lotteries determines the spectrum of risk preferences that will be covered by the risk preference measures. The MPL design accounts for the elicitation of risk preference measures on the whole spectrum, ranging from extreme risk aversion to extreme risk lovingness. In contrast, the outcomes in the OLS-BW design only allow individuals to reveal risk preferences between extreme risk aversion and risk neutrality. Due to the cap at risk neutrality, the OLS-BW experiment does not allow for the differentiation between risk-neutrality and risk-lovingness, or between different degrees of risk-lovingness. In particular when subjects are risk-seeking, crucial information on their risk preferences will be missing.[7] The cap at risk neutrality might further cause a bias toward risk aversion due to the subjects' tendency to categorize themselves somewhere in the middle of an offered range of choices (Likert 1974). In the OLS-BW experiment with its middle options reflecting moderate levels of risk aversion, subjects are likely to pick a more risk averse option than they would in an experiment with a risk neutral option as the mid-point.

Second, the larger the number of choices in an experiment, the larger the number of categories and the finer will be the characterization of risk preferences by the elicited risk preference measures. This implies further more accurate point estimates and an easier convergence when estimating risk preference measures based on structural decision models (Dave et al. 2010; Charness et al. 2013). The MPL design allows for a fine characterization of risk preferences by requesting 102 binary decisions across six price lists; within a single price list of the MPL format alone, respondents' risk preferences are categorized into 16 different Certain Equivalents (CE) and risk preference levels. In contrast, the single decision in the OLS-BW design categorizes respondents merely into six risk preference categories leading to a rough measure of risk preferences.

Third, the design of the lotteries determines for which risk environment preferences will be elicited. While conventional EUT is built on the assumption that risk preferences are an underlying individual trait (von Neumann and Morgenstern 1953), there is mount-

[7]This concern has been considered as negligible as subjects are commonly assumed to be risk averse (Charness et al. 2013). In the literature there is however increasing evidence for the prevalence of risk lovingness, also in developing countries (for example Henrich and McElrath 2002; Maertens et al. 2014; Vieider et al. 2015).

ing evidence that subjects' risk attitudes depend on the risk environment (Kahneman and Tversky 1979). Depending on the probabilities and the framing of a lottery, the elicited risk preferences might then only be valid for the respective risk environment. For example, subjects' choices between lotteries with a high probability risk might only reflect the subjects' risk preferences in a high risk environment and not necessarily be transferable to a low or medium risk environment. The MPL design generates risk preference measures in six different risk environments that are characterized by different probabilities (50%, 25% and 75%) and positive and negative outcomes.[8] As the OLS-BW experiment is restricted to the winning environment with medium-level risk (50%), the elicited OLS-BW risk preference measure might not be valid for a high or low risk environment, or within a losing frame.

Implications on the Noise in the Risk Preference Measures We further identify three design features that have implications on the complexity of the experiment and potentially cause noise in the risk preference measures to be elicited. First, noise might be affected by the number of alternative options in a single decision. The more options, the harder it is for individuals to simultaneously evaluate the pros and cons of each of the different alternatives and to finally make the decision (Scheibehenne et al. 2010). In the OLS-BW design, respondents need to choose between six alternatives in a single decision, which might cause confusion and lead to noise in the risk preference measures. This is less of a problem with the MPL design where respondents can only choose between two options once at a time.

Furthermore, an experiment with different risk environments might come at the cost of more mistakes in the choices and a larger noise in the risk preference measure. In particular, when respondents are not able to realize the differences between the different risk environments, they will not make decisions reflecting their actual preferences in the respective environment. The different risk environments in the MPL design might cause confusion among subjects' and noise in the risk preference measures, while the single risk environment in the OLS-BW design avoids this kind of mistake.

Finally, the number of decisions might affect the level of noise in the risk preferences. On the one hand, with more decisions, subjects get the chance to learn and improve their responses through the course of the experiment, which would then result in less noisy measures (Hey 2001). On the other hand, the more decisions, the longer the experiment

[8]During the pilot, we tested price lists with the probabilities of 12.5% and 87.5%. We found the marginal change in utility of probabilities above 75% and below 25% approached zero. Hence, we did not include the probabilities at the extremes of the distribution in our final experimental design.

takes and the higher the likelihood that subjects will lose concentration and make choices that deviate from their true preferences (Hey and Orme 1994). It is not *a priori* obvious which of these effects will dominate and whether the MPL design with a large number of choices or the OLS-BW design with a single choice will generate lower levels of noise in the risk preference measures.

Independent from the impacts on the levels of noise, a larger number of decisions has the advantage to allow for the direct identification of noise. Only with an experiment where subjects make more than one decision is it possible to detect inconsistencies between choices and to directly identify noise (see detailed explanation in section 3.3.1). With the elicited OLS-BW experiment, it remains unclear to what extent subjects really understood the task and it is impossible to directly investigate inconsistent choices and its effects on the risk preference measures.

To conclude, the MPL design outperforms the OLS-BW design in terms of its prerequisites to be able to generate undistorted, precise and comprehensive measures of risk preferences: the MPL design covers the full range of risk preferences and its large number of decisions enables a much finer characterization than the OLS-BW design. In addition, the MPL design allows for the elicitation of risk preference measures in different environments, including losing and winning lotteries and different risk environments. In terms of their potential to generate a reliable measure of risk preferences, both designs have their shortcomings: while the OLS-BW design does not allow for learning effects and might lead to problems when subjects simultaneously choose among several lottery options, the MPL design potentially causes noise due to fatigue in the course of the experiment and confusion between different risk environments. An advantage of the MPL design is that it allows for the direct identification of noise in form of inconsistent choices. It is essential to understand the different design features and its implications when comparing and interpreting the measures collected from our sample.

3.2.2 Implementation

Recruitment We recruited 875 individuals who were randomly selected for the experiments. During the recruitment phase, the participants were informed that they were eligible to participate in an experiment and a survey. Our sample consists of rural farmers. The average age of the sample is 43 and almost half (44%) are women (see table 3.3). The farmers are characterized by low education levels. More than half (58%) of the sample is illiterate and less than 20% have more than primary education. Furthermore, 10 enumerators were recruited to conduct the experiments.

Protocol Before the experiment, all enumerators received a training in basic risk theory and practiced the games extensively to ensure data quality. Before starting the actual experiment, the experiments were centrally explained. The enumerators illustrated the probabilities in front of all participants by showing yellow and blue-colored tokens for winning and losing. The monetary incentives were demonstrated by putting the respective money value in front of the group. In the explanation of the price lists, we used a price list with a probability of 67.5% that was not used in the actual experiment to avoid any framing effects. The enumerators explained further the rationality in the MPL behind starting with the lottery as the preferred choice and switching once to the safe amount within a given price list (see explanation of rational choices in the MPL in section 3.3.1).[9] After the general explanations, subjects were asked whether they fully understood the instructions. If participants said they did not understand or if there were any remaining questions, the procedure was explained again in more detail.

Once common understanding among the participants could be assumed, the first game was played. First, the enumerators explained and illustrated centrally in front of all participants the probability distribution of the lottery and the outcomes of the first game. Then, the enumerators went to the table of each respondent and visualized again the outcomes and probabilities with a sheet of paper (see figures 3.2 and 3.3) before asking the respondent privately for their preferred choice. If there were any questions remaining, the enumerators clarified individually the open issues at the table. In case a respondent had the intention to switch multiple times within a list of the MPL experiment, the enumerator explained directly to the respondent the rationality behind switching once; if the participant still insisted on switching, the enumerator wrote down the choice. If a respondent wanted to later change a previous decision, the enumerator went back to the respective question and recorded his or her request. After all participants finished their decisions in the first game, the second game was again centrally explained before the respondents were individually asked to make their decisions. This procedure was repeated for all price lists and the OLS-BW. At no time during the experiments were participants allowed to talk to each other. Throughout the session, we repeatedly explained that each single decision mattered for the final payout because at the end of the whole session only one of the decisions would be drawn.

[9]We identify respondents who switched several times within a price list as violators of first-order stochastic dominance (see definition in 3.3.1).

Payment We made a base payment of 50 ETB as an incentive to all participants independent of their decisions and outcomes. In case a lottery with the possibility of winning was drawn, the participant could additionally win up to 160 ETB depending on his or her decision and the outcome of the lottery; in case a lottery with the possibility of losing 160 ETB was drawn, the respondent was additionally supplied with 160 ETB. The overall incentives in the games reflect the daily wage for unskilled labor at the time of the experiment, which ranges between 50 and 150 ETB.

The payout was determined for each participant separately after all decisions were taken. This was conducted through a random lottery incentive mechanism in order to avoid wealth or portfolio effects when subjects make multiple choices.[10] Numbered tokens were randomly drawn for each participant at his or her table. First, there was a random draw between the risk preference and the risk sharing game. In case the risk preference game was drawn, one of the six price lists or the OLS-BW was drawn. In case one of the price lists was selected, an additional draw determined which of the 17 rows of the price list would be played. Finally, winning or losing was determined by drawing a token in the case the participant had chosen a lottery. After the determination of the payout, the respondent received the cash. The hand-over of the cash as well as the drawings for the determination of the payout were done in complete privacy and each individual could observe the draws of the tokens determining his or her payout. The subjects were informed about the details of the payment process before the experiment.

3.3 Measuring Risk Preferences and Noise

3.3.1 Non-parametric Measures

Based on the choices in the OLS-BW and the MPL experiments, we directly calculate non-parametric measures of risk preferences for each individual. The direct identification of noise is only feasible with the MPL as it requires a comparison of at least two choices per individual to identify inconsistent choices. Hence we are not able to compare non-parametric noise between the OLS-BW and the MPL experiment and only analyze non-parametric noise in the MPL choices.

[10] The random lottery incentive mechanism introduces a compound lottery, assuming that subjects consider each choice as equally attractive as the compound lottery that would result when multiplying the probabilities and the single choices (Schoemaker 1982). In a widely cited study, Starmer and Sugden (1991) show that the random lottery incentive mechanism does not create any significant bias. However, there is some evidence (Harrison et al. 2012) showing that people are not indifferent between paying out a single lottery and a randomization of the lottery. If this effect dominates, the random incentive design would create a larger bias in the MPL choices than in the OLS-BW choices.

Direct Measures of Risk Preferences

OLS-BW experiment We infer the level of risk aversion in the OLS-BW game directly for each of the six choices under the assumption of constant relative risk aversion (*crra*) (see table 3.1). The utility U under *crra* is formalized as follows:

$$U(k) = \begin{cases} \frac{k^{(1-r)}}{1-r} & \text{if } r \neq 1 \\ \ln(k) & \text{if } r = 1 \end{cases} \tag{3.1}$$

with k as the outcome, including the experimental endowment and the price of the lottery, and r as the relative risk aversion coefficient; $r > 0$ stands for risk aversion, $r = 0$ reflects risk neutrality and $r < 0$ risk lovingness.[11] By comparing each gamble to its adjacent gambles, we calculate the value of r that generates the same utility level for both payouts and identify a risk aversion range for each of the six OLS-BW choices.[12] In this way, the choice of each individual is directly interpreted as a specific level of risk aversion, ranging from extreme risk aversion (choice 1) to risk neutrality (choice 6).

MPL experiment To get a direct measure of risk preferences from the choices in the MPL experiment, we exploit the respondent's switching point between the lottery and the safe amount within a price list. With increasingly safe amounts in a price list, rational respondents change the decision at some point and prefer the safe amount over the lottery. For example in the winning price list with a chance of 50% (see figure 3.3), respondents are likely to start with preferring the lottery of winning 160 ETB over a safe amount of zero in the first row, and switch at some point to the safe amount and finally prefer 160 ETB for sure over the lottery in the very last row. The switching point determines the CE, which is simply the arithmetic mean of the smallest safe amount a subject had still preferred the lottery and the subsequent safe amount on the list, when the subject preferred (for the first time) the safe amount. The CE covers the whole range of risk preferences, from extreme risk aversion to extreme risk lovingness. We calculate the CEs for each individual in each of the six price lists, reflecting the individual's risk preferences in six different risk environments.

[11] The coefficient of relative risk aversion is defined as the Arrow-Pratt measure (Pratt 1964) as follows: $\frac{-kU''(k)}{U'(k)}$.

[12] For example, a person choosing gamble 3, would have a coefficient of relative risk aversion in the range 1.74-0.81; a person with r=1.74 would be just indifferent between gambles 2 and 3, and a person with r=0.81 is just indifferent between gambles 3 and 4 (see table 3.1).

Direct Identification of Noise

We further investigate choices that are inconsistent and cannot be explained by any rational decision-making theory. By comparing a respondent's choices within the MPL, we identify inconsistent choices by the degree of how serious they violate stochastic dominance:

Violations of First-order Stochastic Dominance (FoSD) refer to dominated and intransitive choices within a given price list. One way to violate FoSD in our MPL design is to choose a strictly dominated option at the beginning or at the end of the row, for example preferring the lottery with a 50% chance of winning 160 ETB over a safe amount of 160 ETB. This violation refers to subjects who never switch (and hence either choose a dominated option at the beginning or at the end of a price list), as well as backward switchers (who choose both at the beginning and the end of a price list a dominated option). Another way of violating FoSD is to switch several times within a price list, reflecting intransitive choices.[13] FoSD violations are considered to be a clear indicator of poor understanding, as subjects are able to see and easily compare previous decisions within a list when making their choice.

Violations of Second-order Stochastic Dominance (SoSD) consist of intransitive choices between several price lists. For example, a subject who prefers the lottery over the safe amount of 100 ETB for sure in the price list with a 50% chance of winning, should also prefer the lottery when choosing between 100 ETB for sure and the lottery in the price list with a 75% chance to win. A subject would violate SoSD with preferring the lottery in the price list with the lower winning chance, while preferring the safe amount in the price list with the higher winning chance. As respondents are not able to directly compare their choices in previous price lists, SoSD violations might not necessarily reflect a serious lack of understanding.

3.3.2 Parametric Measures and Economic Models

We further investigate both elicitation methods by parametrically estimating risk preferences based on standard economic models. The parametric risk preference measures reflect the average risk preferences in our sample.[14] In comparison to analyzing non-

[13] For example, a subject preferred the lottery up to a safe amount of ETB 90, switched then and indicated a preference for the safe amount of ETB 100 over the lottery, and switched later again with a preference for the (same) lottery over a safe amount of ETB 110.

[14] The parametric estimation of individual-level risk preferences is not feasible with our data as it would require a larger number of decisions per individual.

parametric measures, this offers two advantages: It allows us to analyze risk preferences based on economic theory and as such to identify the prevalence of non-*crra* preferences and subjective probability weighting in our sample. Furthermore, the parametric estimations allow us to estimate parameters of stochastic noise in both OLS-BW and MPL choices, controlling for individual heterogeneity in the unobservables that otherwise potentially distorts the measures of risk preferences (Harrison and Rutström 2008).

Expected Utility Theory (EUT)

To analyze basic risk preference measures, we first estimate risk preference and noise parameters based on EUT (von Neumann and Morgenstern 1953), the classical theory of behavior and decision-making. It assumes that subjects make choices depending on the expected utility from the lottery, which is derived from the expected value of the outcomes of the lottery, their initial wealth and their preferences regarding risk.

EUT with *crra* preferences We start with the assumption of *crra* preferences (equation 3.1), with risk preferences r being independent of the level of wealth of the individual and the price of the lottery k. Under EUT, the decision-maker is assumed to weight each possible outcome $k_c \in \{1,...,K\}$ in choice $c \in \{1,...,C\}$ using the objective probability p_{k_c}, which is associated with the outcome. The expected utility from choice c (EU_c) is then equal to the sum of the probability weighted utility for each outcome U_{k_c}:

$$EU_c = \sum_{k_c=1}^{K} p_{k_c} U_{k_c}. \tag{3.2}$$

In the MPL procedure, respondents choose in each question $q \in \{1,...,17\}$ within each price list $l \in \{1,...,6\}$ between a risky option and a safe amount. To make their decisions, the participants compare repeatedly between the two options $c \in \{1,2\}$ and the expected values EU_1 and EU_2.

To fit this model to the data and to estimate the risk preference parameter r, we use a structural model combined with a maximum likelihood estimation technique, following the approach developed by Camerer and Ho (1994) and elaborated by Harrison and Rutström (2008). With the MPL choices, first, the expected utility $EU_c^{l,q}$ from each potential choice c in each question q and each price list l is calculated according to equation 3.2.

Then, a latent index $\triangle EU_c^{l,q}$ is calculated by linking latent risk preferences to choice probabilities as follows:

$$\triangle EU_c^{l,q} = \frac{\exp(EU_c^{l,q}/\mu)}{\exp(EU_1^{l,q}/\mu) + \exp(EU_2^{l,q}/\mu)} \tag{3.3}$$

with the error term μ capturing the randomness in the choices. The smaller μ, the smaller the errors in the decision-making process and the closer the observed choices to the true, deterministic preferences. The index $\triangle EU_c^{l,q}$ is in the form of a probability between 0 and 1; thus it can be directly linked to the observed choices c. It can be interpreted as the probability of a subject choosing c in question q of price list l.

The inclusion of a stochastic error term μ allows observed choices to deviate from true preferences and provides information about the extent of error predictions (Wilcox 2008). Stochastic errors may arise due to different reasons: individuals might face comprehension problems of the experimental task, they might be careless or get tired during the course of the experiment, or they might not be aware of the own utility function when making their choices (Hey 1995; Loomes 2005). This further relates back to the design of the experiment. Stochastic noise might also be influenced by the quality of the explanations by the enumerators and by disturbing factors during a session. A different reason for the occurrence of stochastic noise might be the misspecification of the model such that the model does not well describe the decisions of the sample.

The log-likelihood of the observed MPL choices of all participants N in all 17 questions q in all six price lists l is defined as follows:

$$\ln L^{MPL}(r, \mu; \mathbf{y}) = \sum_{i=1}^{N} \sum_{l=1}^{6} \sum_{q=1}^{17} \ln(\triangle EU_{y_i^{l,q}}^{l,q}) \tag{3.4}$$

with $\mathbf{y}$ is the vector of observed choices in all questions and price lists of individual i.

In the OLS-BW procedure, the respondents make only a single decision between six choices $c \in \{1, 6\}$ and its expected utilities (EU_1, ..., EU_6). We calculate the expected utility EU_c for each choice $c = \{1, ..., 6\}$ according to equation 3.2, and then derive the latent index $\triangle EU_c$ for each choice, linking latent risk preferences to the choice probabilities, as follows:

$$\triangle EU_c = \frac{\exp(EU_c/\mu)}{\sum_{j=1}^{6}\exp(EU_j/\mu)}. \tag{3.5}$$

Based on the choice probabilities, the log-likelihood of the observed choices of all participants in the OLS-BW lottery is formulated as follows:

$$\ln L^{OLS-BW}(r,\mu;\mathbf{y}) = \sum_{i=1}^{N}\ln(\triangle EU_i) \tag{3.6}$$

with **y** is the vector of the OLS-BW choice of all individuals.

We maximize separately the log-likelihood of the MPL and the OLS-BW choices in equations 3.4 and 3.6, and derive estimates on the risk preference and stochastic noise parameters, by using conventional numerical optimization algorithms[15] and the statistical software *Stata* (version 13.1).[16] The standard errors of the maximum likelihood estimations are clustered at the subject-level, allowing for the possibility of correlation between choices.[17]

In addition, we assess the OLS-BW and the MPL choices simultaneously to account for the mutual influence of both experiments (following Andersen et al. 2008 and Andersen et al. 2014). We control for choices, which were conditionally made on knowing the choices in earlier tasks, and for error propagation effects between the experiments by maximizing the joint likelihood of the OLS-BW and the MPL choices as follows:

$$\ln L(r,\mu;\mathbf{y}) = \ln L^{OLS-BW} + \ln L^{MPL}. \tag{3.7}$$

The simultaneous estimation of OLS-BW and MPL choices allows us to test for the difference in risk preference and stochastic noise parameters between both experiments (Harrison and Lau 2014).

[15] We used *Stata's* Newton-Raphson, Davidon-Fletcher-Powell and Broyden-Fletcher-Goldfarb-Shanno algorithms.

[16] A detailed explanation on how to run maximum likelihood estimations in *Stata* can be found in Harrison (2008).

[17] Clustering is commonly used in the literature and allows for heteroskedasticity between and within clusters, as well as autocorrelation within clusters (Andersen et al. 2008).

EUT with *expo-power* preferences To analyze the interactions between risk preferences and wealth, we relax the assumption of *crra* preferences and adopt the *expo-power* (*ep*) utility, which was originally proposed by (Saha 1993). Following Holt and Laury (2002), the *ep* function is defined as

$$U(x) = \frac{(1 - \exp(-\alpha_{ep} k^{1-r_{ep}}))}{\alpha_{ep}} \tag{3.8}$$

with the parameters α_{ep} and r_{ep} to be estimated. The risk preference measure under *ep* preferences can be calculated as follows:

$$r = r_{ep} + \alpha_{ep}(1 - r_{ep})k^{1-r_{ep}}. \tag{3.9}$$

The *ep* function incorporates both increasing relative risk aversion (*irra*) and decreasing relative risk aversion (*drra*), which is determined by the parameter α_{ep}. If $\alpha_{ep} = 0$, risk aversion is constant with wealth, while $\alpha_{ep} > 0$ and $\alpha_{ep} < 0$ stand respectively for increasing and decreasing risk aversion with wealth. r_{ep} determines whether risk aversion is defined over absolute ($r_{ep} = 0$) or relative ($r_{ep} \neq 0$) wealth. Hence, the *ep* preference function nests both *crra* preferences ($\alpha_{ep} \to 0$) and constant absolute risk aversion (*cara*) ($r_{ep} \to 0$).

Similar to the procedure described above with *crra* preferences, the expected utility with *ep* preferences in equation 3.8 is evaluated according to equation 3.2, and the model is fit to the data by using the structural models in equations 3.3 and 3.5, and the log-likelihood functions in equations 3.4 and 3.6 for MPL and OLS-BW choices respectively.

Rank-dependent Utility (RDU) Theory

We further test whether individuals evaluate low-probability risks differently than large-probability risks and assume RDU theory (Harless and Camerer 1994). The RDU model allows subjects to weight probabilities associated with final outcomes in a non-linear manner. RDU consists of two components: the utility function and the probability weighting function. Instead of weighting outcomes with objective probabilities p_{k_c} (as under EUT), a probability weighting function is used to account for the subjective importance of different outcomes. The rank-dependent utility from choice c is defined as follows:

$$RDU_c = \sum_{k_c=1}^{K} w_{k_c} U_{k_c} \tag{3.10}$$

with the outcomes ranked from worst (U_{1_c}) to the best (U_{K_c}). The cumulative probabilities w_{k_c} are defined as:

$$w_{k_c} = \begin{cases} \omega(p_{1_c}) & \text{for} \quad k_c = 1 \\ \omega(p_{1_c} + ... + p_{k_c}) - \omega(p_{1_c} + ... + p_{(k-1)_c}) & \text{for} \quad k_c > 1 \end{cases} \tag{3.11}$$

with $\omega(p)$ as the commonly used Tversky-Kahneman (1992) probability weighting function:

$$\omega^{TK}(p) = p^{\gamma}/[p^{\gamma} + (1-p)^{\gamma}]^{1/\gamma} \tag{3.12}$$

for $0 < p < 1$, with the well-defined endpoints $\omega(p) = 0$ for $p = 0$, and $\omega(p) = 1$ for $p = 1$.[18] Subjective probability weighting is imposed by $\gamma \neq 1$. For the case of $0 < \gamma < 1$, the probability weighting function is "inverse s-shaped" and extremely unlikely outcomes (with very low probabilities) are overweighted up to a cross-over-point at $\omega(p) = p$, after that very likely outcomes (with high probabilities) are underweighted. For the case of $\gamma > 1$, the probability weighting function is "s-shaped" indicating the opposite behavior. The probability weighting function can be combined with any kind of utility function, such as the *crra* and the *ep* functions.

The RDU model requires choices in different risk environments, to be more specific, lotteries with different probabilities are needed. Hence, we can only fit the RDU model to our MPL choices and not to the OLS-BW choices. We estimate risk preference under RDU with the parameter for subjective probability weighting γ by using equations 3.10 and 3.11, combined with the Tversky-Kahneman probability weighting function (equation 3.12) and both *crra* and *ep* utility functions (equations 3.1 and 3.8), with and without allowing for stochastic errors. The log-likelihood functions are optimized according to equation 3.4 using the MPL choices to obtain structural maximum likelihood estimates of the utility function parameters and the probability weighting parameter γ.

Heterogeneity of Stochastic Noise

To elaborate the drivers of stochastic noise in the risk preference measures, we further allow stochastic noise to be a linear function of observable factors (following Andersen et al. 2008). We allow the estimated stochastic noise parameter $\hat{\mu}$ to depend on individual

[18] We applied also the two-parametric Prelec function, but not all models converged with our data; for the models that converged, results were similar.

characteristics and on features that characterize how the experiments were implemented. This is formulated as follows:

$$\hat{\mu} = \hat{\mu_0} + \sum_{c=1}^{4} (\hat{\mu}_{ind_c} * ind_c) + \sum_{d=1}^{9} (\hat{\mu}_{enu_d} * enu_d) + (\hat{\mu}_{time} * time) \tag{3.13}$$

where $\hat{\mu_0}$ is the estimate of the constant; *ind* is a vector of the individual characteristics, including age, a gender dummy, the number of livestock[19] as a proxy for wealth, and a dummy indicating whether the participant is literate; *enu* is a vector with dummies for the enumerators who requested the individual's choices; $time$ is a dummy variable indicating whether the experiment was conducted in the morning or in the afternoon. The different $\hat{\mu}_{ind_c}$ and $\hat{\mu}_{enu_d}$ estimates, and the $\hat{\mu}_{time}$ estimate, show the differences in stochastic noise for the respective control variables.

The log-likelihoods of the observed MPL and OLS-BW choices of all participants are then maximized as follows:

$$\ln L^{MPL}(r, \mu; \mathbf{y}, \mathbf{X}) = \sum_{i=1}^{N} \sum_{l=1}^{6} \sum_{q=1}^{17} \ln(\triangle EU^{l,q}_{y_i^{l,q}}) \tag{3.14}$$

$$\ln L^{OLS-BW}(r, \mu; \mathbf{y}, \mathbf{X}) = \sum_{i=1}^{N} \ln(\triangle EU_i). \tag{3.15}$$

with **X** is the vector that includes individual- and session-specific characteristics.

3.4 Empirical Analysis

In this section, we empirically investigate the OLS-BW and the MPL experiment by exploiting the choices from our sample of Ethiopian farmers. In the following, we compare both elicitation methods in their ability to measure risk attitudes and to minimize noise.

3.4.1 Investigation of Risk Preference Measures

To investigate levels of risk preferences and how they change in different risk environments, we first examine non-parametric measures of risk preferences before parametrically estimating the risk preferences of our sample. The parametric estimation allows us

[19] We transfered the number of livestock in tropical livestock units according to the FAO scale (1982).

to disentangle stochastic noise and to investigate in more detail the interactions between risk preferences and wealth as well as between different risk environments.

Non-Parametric Measures We calculate non-parametric measures of risk preferences for the OLS-BW and the MPL choices as described in section 3.3.1. The non-parametric measures from the OLS-BW and the MPL choices do not give a consistent picture of our sample's risk preferences. The MPL measure illustrates our sample as being risk loving or risk neutral, while the OLS-BW measure identifies them as moderately risk averse: All CEs in the price lists are either above (=risk lovingness) or equal (=risk neutrality) to the EV of the respective lotteries (see table 3.4), while the average choice in the OLS-BW lottery is between the two middle categories reflecting moderate risk aversion (see figure 3.4). This difference can be explained by the cap at risk neutrality in the design of the OLS-BW experiment. A large proportion - over one-third - of our sample chose the highest possible category of risk neutrality/lovingness in the OLS-BW task. Without the cap, individuals might have chosen a lottery reflecting higher levels of risk lovingness. This is evident in the choices in the price list that is most comparable to the OLS-BW (with a winning chance of 50%), where over two-thirds of our sample made risk neutral or risk loving choices (see figures 3.4 and 3.5a).[20] The individuals also rated themselves as risk loving when directly asked about their willingness-to-take risk in the survey (see table 3.5). Given the heavy distortion toward risk aversion in the OLS-BW measure, the MPL measure seems to be preferable when relying on non-parametric methods to describe a sample's risk preferences.

Furthermore, we analyze different dimensions in the non-parametric risk preference measures. As, by design, the OLS-BW experiment covers only a single risk environment (characterized by medium-level risk and the winning domain), we examine the differences in the non-parametric risk preference measures from the MPL experiment. More specifically, we investigate whether and to what extent the non-parametric measures from the six price lists differ, to examine the necessity of an experimental design that covers different risk environments for our sample. We find the non-parametric measures of the price lists reveal slightly different risk preference levels (see table 3.4). Subjects show different levels of risk lovingness when the chances of winning are low or medium (25% or 50%) or the chances of losing are medium or high (50% or 75%) as indicated by the significantly positive differences between the CEs and the EVs; the respondents are risk-neutral within environments where the chances of winning are high (75%) or the chances

[20]Specifically, 31.66% of the individuals choose the risk neutral/risk loving category in the OLS-BW experiment. In the most comparable price list, 27% and 35% of the individuals make choices revealing risk neutral and risk loving preferences respectively.

of losing are low (25%). We find that this difference is mostly driven by violators of SoSD who had problems with staying consistent over the different price lists. Restricting our sample to subjects who did not violate SoSD dominance still reveals different levels of risk lovingness in the different price lists. We do not find any significant differences in risk preferences between the winning and losing environment. To sum, we find the non-parametric MPL measures are able to show nuances of risk preferences in the different risk environments and that the winning or losing domain is less important with our sample. This implies that the non-parametric OLS-BW measure might in principle be transferable to a losing environment and even roughly reflect the risk preferences in other risk environments, if it was not capped at risk neutrality.

Parametric Measures We investigate the risk preference measures in greater detail by parametrically estimating risk preferences as explained in section 3.3.2. We first estimate the risk preferences based on EUT-*crra* preferences both separately and jointly for the OLS-BW and MPL choices according to equations 3.4 and 3.6, and equation 3.7. Both OLS-BW and MPL risk preference parameters reveal risk lovingness within our sample: The estimated risk preference coefficients are negative and significant at a 1% significance level (see tables 3.6 and 3.7).[21] Although there are slight differences in the levels of risk lovingness, both experiments give a similar picture on individuals risk preferences that is in line with the results from the non-parametric MPL measures and the survey responses. This implies that the parametric estimation seems to correct for the distortion caused by the design of the OLS-BW experiment.

We further investigate parametrically the need for an elicitation design that incorporates different dimensions of risk preferences with our sample. We first test whether non-*crra* preferences are relevant, estimating risk preferences based on the EUT model with *ep* preferences.[22] We find the prevalence of *drra* preferences with our sample of Ethiopian farmers both in the MPL and OLS-BW choices (see table 3.8). This is evident in significantly negative α_{ep} coefficients, and r_{ep} coefficients that are significantly different from zero (see columns 1-3). In other words, risk aversion is decreasing (and consequently risk lovingness is increasing) with increasing outcomes relative to wealth. We further experience that the estimation of the OLS-BW choices is only robust for the EUT-*ep*

[21] When estimating the EUT-*crra* model without noise, the MPL coefficient is significantly positive and different from the OLS-BW risk preference parameter; however, this result is not reliable as of the positive and significant μ parameter in the model with noise, thus indicating that noise should be included to draw any conclusions.

[22] The joint estimation of the EUT-*ep* model based on OLS-BW and MPL choices does not converge with our data; hence, we maximize the log likelihood functions for the OLS-BW and the MPL choices separately according to equations 3.4 and 3.6; this implies that the size of the estimated coefficients is not necessarily comparable.

model without noise, and it is not feasible to simultaneously estimate *ep*-coefficients and stochastic noise without generating a huge noise parameter (see column 4).
Furthermore, we analyze parametrically the interaction between risk preferences and different risk environments, and test for the prevalence of subjective probability weighting. By design, this is only feasible with the MPL, which covers different risk environments. Using the MPL choices, we extent the analysis to RDU models and run several specifications (RDU-*crra*, RDU-*ep*, with and without noise). In all specifications, we find our sample systematically overweights low probability risks and underweights high probability risks. This is evident in the significant γ coefficients which are smaller than one in all specifications, rejecting the hypothesis of linear probability weighting (γ=1) (see table 3.9).

To conclude, both parametric OLS-BW and MPL risk preference measures reveal the prevalence of risk lovingness and *drra* preferences with our sample of Ethiopian farmers. The OLS-BW experiment seems to be sufficient when risk preferences are parametrically estimated, while we find the non-parametric measure to be heavily distorted toward risk aversion. When investigating risk preferences in different risk environments, the MPL outperforms the OLS-BW experiment. It is able to identify nuances of risk lovingness in the different risk environments and to reveal an overweighting of low probability risks and an underweighting of high probability risks from our sample.

3.4.2 Investigation of Noise in the Risk Preference Measures

We further investigate the extent of noise to explore the reliability of the OLS-BW and MPL risk preference measures and to examine whether the experiments are too complex for our low-educated sample of Ethiopian farmers. Even if the risk preference measures reflect individual risk attitudes, they might include substantial noise and hence not necessarily be reliable. To assess the extent of non-parametric noise, we identify violators of first- and second-order stochastic dominance in the MPL experiment. We further test whether the estimated risk preference parameters are robust to the exclusion of violators of stochastic dominance. To compare the reliability of the OLS-BW and the MPL risk preference measures, we investigate the estimated parameters of stochastic noise based on the choices from our sample and elaborate potential drivers, including individual characteristics and the influence of the enumerators and the daytime of the session.

Direct Identification of Noise When directly identifying inconsistent choices in the MPL experiment (see explanation in section 3.3.1), we find moderate levels of violations of stochastic dominance. FoSD was violated by 8.69% (or 76) of the participants, who mostly violated stochastic dominance in only one or two of the price lists.[23][24] Almost all of them (73 participants) violated FoSD by never-switching between the lottery and the safe amount. In this way, they made inconsistent choices either at the beginning or at the end of the rows of a price list. Less than 1% (2 participants) violate both at the beginning and the end of the rows by backward-switching. Few subjects violate FoSD by multiple switching within a price list, with less than 1% (or 4 participants) switching several times.[25] The SoSD violations are also at a moderate level, but as expected higher than the FoSD violations. We find 15.09% (or 132 subjects) of our sample violate SoSD at least once. FoSD and SoSD violations seem not to be closely related, as indicated by only 3.3% of the sample violating both FoSD and SoSD. To elaborate whether stochastic dominance violations distort the estimates of the risk preference parameters, we estimate the EUT-*crra* model without violators of stochastic dominance. Table 3.10 shows the estimates when excluding FoSD and SoSD violators. We find that the estimated risk preferences do not change, neither when excluding FoSD violators and SoSD violators, nor when excluding both. The moderate levels of inconsistent choices do not distort our main findings and indicate that the MPL is sufficiently well understood by our low-literate sample.

Parametrically Estimated Stochastic Noise In all previous estimations of the EUT and RDU models, we find evidence for stochastic noise in the OLS-BW and MPL choices as indicated by μ coefficients that are significantly different from zero (see tables 3.6, 3.7, 3.8, 3.9 and 3.10). We are able to compare stochastic noise in the EUT-*crra* specification for which we receive robust estimates of the noise parameters both with the OLS-BW and the MPL choices.[26] The separate and joint estimations of the EUT-*crra* model (see tables 3.6 and 3.7) reveal moderate levels of stochastic noise both in the OLS-BW and the MPL choices, and we cannot conclude that the more comprehensive MPL experiment

[23]Specifically, there were 55 individuals violating FoSD in one, and 11 individuals violating in two price lists; Only 4 respondents violated FoSD in more than 3 price lists.

[24]The level of FoSD violations with low-literate samples in developing countries in our MPL format is much lower in comparison to the classic Holt and Laury (2002) MPL design which is found to cause FoSD violation rates between 35% (Humphrey and Verschoor 2004) and over 50% (Jacobson and Petrie 2009; Charness and Viceisza 2012).

[25]Multiple switching is usually considered as a violation of stochastic dominance as it violates the transitivity assumption. However, multiple switching could also result from genuine indifference of the subject between alternative choices (Andersen et al. 2006).

[26]The level of noise in the estimations of the EUT-*ep* model are at such a high level with the OLS-BW choices such that we rather assume this is due to an insufficient number of observations in the OLS-BW experiment rather than a lack of misunderstanding or other subject-related reasons (see column 4, table 3.7).

causes higher levels of noise than the simple OLS-BW.

To better understand potential drivers of stochastic noise in the choices of our sample, we finally add different sets of covariates to explain stochastic noise within the EUT-*crra* model as described in section 3.3.2.[27] First, we exploit information on individual characteristics. To our surprise, we do not find that any of the individual characteristics explains stochastic noise, neither in the OLS-BW nor in the MPL choices (see table 3.11). Differences in stochastic noise within our sample cannot be explained by differences in literacy, age, gender or livestock wealth. Furthermore, we test whether stochastic noise is influenced by the way the experiments were implemented. We first investigate whether the daytime of the experiment, which might influence the respondents' ability to concentrate, affects levels of stochastic noise. We find a significant effect of the daytime on stochastic noise in the MPL choices. The level of stochastic noise is significantly larger in the choices made during the afternoon sessions. This implies that mistakes in the MPL experiment, for instance due to fatigue, are more common during the afternoon. In contrast, the choices in the OLS-BW experiment seem to be unaffected by the timing of the experiment. To elaborate the influence of the enumerators on the levels of stochastic noise, we finally add enumerator dummies to stochastic noise in our structural model.[28] Enumerators might be different in their ability to ask participants for their choices and to clarify questions of the participants, which might affect the level of stochastic noise. In both experiments, we find that enumerators significantly affect estimated levels of stochastic noise.[29] There are both enumerators who significantly increase levels of stochastic noise, and enumerators who significantly decrease levels of stochastic noise.

To sum, in both OLS-BW and MPL choices, we find the evidence of stochastic noise. A comparison of noise in the estimations of the EUT-*crra* model shows that we cannot conclude that the MPL experiment causes higher levels of stochastic noise than the OLS-BW experiment. When investigating the drivers of stochastic noise, we find stochastic noise in the MPL risk preference measures is significantly higher when the experiment is conducted during the afternoon. Furthermore, enumerators have a significant influence on the level of stochastic noise in both OLS-BW and MPL experiments. We do not

[27]We are only able to show results from separate estimations as the model does not converge with the joint maximum likelihood when including covariates; this implies that the size of the estimated coefficients is not necessarily comparable.

[28]The daytime dummy and the dummies for the enumerators were added separately, as the model does not converge with the OLS-BW choices when simultaneously controlling for daytime and enumerators.

[29]When controlling for enumerator fixed-effects in estimation of OLS-BW choices, the estimated level of stochastic noise becomes very large; this implies that the estimated size of the effect might not be reliable.

find that noise can be explained by our measures of individual characteristics. The MPL experiment allows us further to identify moderate levels of inconsistent choices.

3.5 Conclusion

While individual risk preferences are key information in the economic analysis of risk management, there is no established method to elicit attitudes toward risk. This paper assesses the suitability of two commonly used methods to elicit risk preferences from low-literate samples in developing countries - the simple OLS-BW and a more comprehensive MPL with a risky option and a safe amount. We investigate risk preference measures and noise from the two experiments empirically by using parametric and non-parametric methods and choices from a sample of farmers in the Ethiopian Tigray region.

We find that both the parametric OLS-BW and the MPL risk preference experiments elicit risk loving preferences with moderate levels of stochastic noise from our sample. When investigating the drivers of stochastic noise, we find that noise in the MPL choices is significantly higher when the game is played during the afternoon (rather than the morning) and that the enumerators significantly influence levels of noise in both experiments. Due to the cap in the range of outcomes, we find the non-parametric OLS-BW risk preference measure heavily distorted toward risk aversion. In contrast to the OLS-BW experiment, the multitude of choices with different outcomes in the MPL experiment allows us to estimate RDU models, revealing an overweighting of low probability risks. The MPL allows us further to identify moderate levels of inconsistent choices. The results indicate that both the MPL and the OLS-BW experiment seem to be well understood by our low-educated sample. Our findings suggest using the simple OLS-BW method when estimating simple risk preferences. They encourage the usage of the MPL procedure when more complex structural models are needed, for instance to analyze subjective probability weighting, when analyzing inconsistent choices and when relying on non-parametric measures to characterize risk preferences.

There is still scope for future research on the elicitation of risk preferences in developing countries which cannot be answered with the data from our experiments. First, the OLS-BW experiment should be adapted and extended to several price lists. On the one hand, this would allow investigating whether the distortion in the non-parametric OLS-BW measure could be easily fixed by an extension of the range of outcomes that includes different degrees of risk-lovingness. On the other hand, inconsistent choices in the OLS-BW risk preference measure could be explored which would open the door for a comparison of noise between the MPL and OLS-BW choices from a non-parametric perspective.

Furthermore, it would be important to better understand how individual skills and abilities affect the elicited risk preference measures. Our analysis of stochastic noise relies on basic information about the education level which can be considered as a rather rough measure of individual skills of our sample of low-educated farmers. A more detailed and precise measure, reflecting for instance the participants' numerical skills and their intellectual abilities, might be able to shed more light on the influence of individual capabilities on the elicited risk preference measures and the understanding of the participants.

In addition to the OLS-BW and the MPL experiments, there are other elicitation approaches used in developing countries that could be evaluated. One other approach is the Gneezy and Potters (1997) design, which offers a single-choice investment task to subjects who are equipped with a basic endowment and then asked how much to invest in a risky lottery. A further method is the willingness-to-pay format (for example Maertens et al. 2014) where participants are given an endowment and then asked for their willingness-to-pay for a range of several lottery options. These two alternative methods are however not sufficient for a parametric estimation of structural decision-making models which implies that comparisons of parametric risk preferences measures are infeasible.

Finally, an important purpose to elicit risk preferences is to achieve individual risk preference measures that are able to explain individual behavior and to predict decision-making. To fully assess the effectiveness of risk preference measures, the explanatory power of individual-level risk preferences measures on behavior and decision-making should be compared. This requires however detailed information on individual's risk preferences and their risk management decisions. In particular, the estimation of risk preference and subjective probability parameters at the individual level poses a big challenge, as it requires a large number of binary decisions between different risky options.

Tables and Figures

Table 3.1: OLS-BW Design

choice	lottery	EV	*crra* parameter	risk aversion class
A	(40, p=0.5; 40)	40	[+∞; 7.51]	extreme
B	(75, p=0.5; 35)	55	[7.51; 1.74]	severe
C	(90, p=0.5; 30)	60	[1.74; 0.81]	intermediate
D	(120, p=0.5; 20)	70	[0.81; 0.32]	moderate
E	(150, p=0.5; 10)	80	[0.32; 0.00]	slight-neutral
F	(160, p=0.5; 0)	80	[0.00; -∞]	neutral-negative

Notes: The table shows each lottery option and its EV with the related risk aversion class based on *crra* preferences. Respondents choose one of the six lotteries as their preferred choice. The lottery values are in ETB.

Table 3.2: MPL Design

price list	lottery	range of safe amount
gain domain		
1	(160; p=0.50; 0)	[0, 10, ..., 160]
2	(160; p=0.25; 0)	[0, 10, ..., 160]
3	(160; p=0.75; 0)	[0, 10, ..., 160]
loss domain		
4	(-160; p=0.50; 0)	[-160, -150, ..., 0]
5	(-160; p=0.25; 0)	[-160, -150, ..., 0]
6	(-160; p=0.75; 0)	[-160, -150, ..., 0]

Notes: In each of the six price lists, respondents repeatedly choose between the lottery and an increasingly safe amount of winning or losing. In each price list, respondents take 17 decisions. The values of the lotteries are in ETB.

Table 3.3: Summary Statistics of Individual Characteristics

	mean	sd	min	max	N
general characteristics					
age	43	13	18	90	861
gender (1=female)	0.44	0.50	0	1	865
literate (1=literate)	0.42	0.49	0	1	865
education (1=higher than primary)	0.19	0.39	0	1	842
livestock wealth					
cows	1	1	0	10	863
bulls	0.44	0.72	0	6	860
oxen	1	0.91	0	4	864
sheep	0.48	1.50	0	15	862
goats	1.10	3	0	30	861
horses	0.01	0.10	0	2	861
camels	0.01	0.10	0	3	861

Notes: The table shows the individual characteristics of our sample of Ethiopian farmers. Not all of the 875 participants answered the survey questions after the experiments.

[ht]

Table 3.4: Summary Statistics of the Price Lists

probability	EV	CE (mean)	CE (sd)	CE-EV	test: EV=CE
gain domain					
p=0.25	40	59.81	40.93	19.81	p<0.01
p=0.5	80	90.62	39.41	10.62	p<0.01
p=0.75	120	119.03	39.22	-0.97	p>0.10
loss domain					
p=0.25	-40	-39.57	36.18	0.43	p>0.10
p=0.5	-80	-69.42	37.42	10.58	p<0.01
p=0.75	-120	-97.86	41.55	22.14	p<0.01

Notes: The table shows for each price list the EV and the mean and standard deviations (st) of the CEs of our sample of Ethiopian farmers. P-values are reported from a t-test on whether EV equals the average CE in a given list (indicating risk neutrality). The values are in ETB.

Table 3.5: Summary Statistics of the Willingness-to-take-Risks Questions

	mean	sd	min	max	N
willingness to take risks ...					
... in general	7.62	1.92	0	10	865
... when deciding on a loan	7.48	2.43	0	10	865
... when buying seedlings	7.47	2.34	0	10	865
... when buying fertilizer	8.67	2.03	0	10	865
... when negotiating over crop prices	6.15	2.79	0	10	865
... when negotiating over share cropping	5.49	3.18	0	10	863

Notes: The table shows the summary statistics of our sample's responses to different questions on willingness-to-take-risks. Each participant received 10 tokens and indicated her level of willingness to-take-risks by putting as many tokens on the table as she thinks her level of risk-taking is, with 10 tokens standing for fully willing to take risks and no token for fully avoiding risks. Not all of the 875 participants in the experiments answered the willingness-to-take risk questions after the experiments.

Table 3.6: Estimated Risk Preferences based on EUT-*crra* Model (Separate Estimation)

	(1)	(2)
model	EUT-*crra*	EUT-*crra*
experiments	*MPL/OLS-BW*	*MPL/OLS-BW*
MPL		
r	0.725***	-0.555***
	(0.000)	(0.000)
μ		3.097***
		(0.000)
observations	89216	89216
OLS-BW		
r	-1.245***	-2.248***
	(0.000)	(0.000)
μ		12.060*
		(0.046)
observations	875	875

Notes: The table shows the results from separate maximum likelihood estimations on the choices from the MPL and the OLS-BW based on the EUT-*crra* model without and with noise. The P-values of the estimated coefficients are based on clustered standarderrors and reported in parentheses. *** p<0.01; ** p<0.05; *p<0.1 significance level.

Table 3.7: Estimated Risk Preferences based on EUT-*crra* Model (Joint Estimation)

	(1)	(2)
model	EUT-*crra*	EUT-*crra*
experiments	*MPL/OLS-BW*	*MPL/OLS-BW*
MPL		
r	0.725***	-0.554***
	(0.000)	(0.000)
μ		3.097***
		(0.000)
OLS-BW		
r	-1.245***	-2.248***
	(0.000)	(0.000)
μ		12.059***
		(0.000)
observations	178466	178466

Notes: The table shows the results from a joint maximum likelihood estimation on the choices from the MPL and the OLS-BW based on the EUT-*crra* model without and with noise. The OLS-BW choices were reweighted to make them comparable. P-values of the estimated coefficients are based on clustered standard errors and reported in parentheses. *** p<0.01; ** p<0.05; *p<0.1 significance level.

Table 3.8: Estimated Risk Preferences based on EUT-*ep* Model

	(1)	(2)	(3)	(4)
model	EUT-*ep*	EUT-*ep*	EUT-*ep*	EUT-*ep*
experiment	*MPL*	*MPL*	*OLS-BW*	*OLS-BW*
r_{ep}	0.716***	0.935***	1.311***	9.044***
	(0.000)	(0.000)	(0.000)	(0.000)
α_{ep}	-0.350***	-16.643***	-14.543*	33.852*
	(0.000)	(0.000)	(0.010)	(0.043)
μ		6.478***		405.555***
		(0.000)		(0.000)
observations	89216	89216	875	875

Notes: The table shows separate maximum likelihood estimates of the MPL and OLS-BW choices based on an EUT-*ep* specification without and with noise. P-values of the estimated coefficients are based on clustered standard errors and reported in parentheses. *** p<0.01; ** p<0.05; *p<0.1 significance level.

Table 3.9: Estimated Risk Preferences based on RDU Models

	(1)	(2)	(3)	(4)
model	RDU-*crra*	RDU-*crra*	RDU-*ep*	RDU-*ep*
experiment	*MPL*	*MPL*	*MPL*	*MPL*
r / r_{ep}	0.716***	-0.466***	0.711***	0.913***
	(0.000)	(0.000)	(0.000)	(0.000)
γ	0.857***	0.867***	0.823***	0.865***
	(0.000)	(0.000)	(0.000)	(0.000)
α_{ep}			-0.348***	-10.259***
			(0.000)	(0.000)
μ		2.904***		4.666***
		(0.000)		(0.000)
observations	89216	89216	89216	89216

Notes: The table shows maximum likelihood estimates of the MPL choices based on RDU-*crra* and RDU-*ep* specifications without and with noise. P-values of the estimated coefficients are based on clustered standard errors and reported in parentheses. *** p<0.01; ** p<0.05; *p<0.1 significance level.

Table 3.10: Estimated Risk Preferences based on EUT-*crra* Model without Violators

	(1)	(2)	(3)	(4)	(5)	(6)
model	EUT-*crra*	EUT-*crra*	EUT-*crra*	EUT-*crra*	EUT-*crra*	EUT-*crra*
experiment	*MPL/OLS-BW*	*MPL/OLS-BW*	*MPL/OLS-BW*	*MPL/OLS-BW*	*MPL/OLS-BW*	*MPL/OLS-BW*
without violators	FoSD/SoSD	FoSD	SoSD	FoSD/SoSD	FoSD	SoSD
MPL						
r	0.696***	0.705***	0.708***	-0.477***	-0.477***	-0.523***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
μ				2.843***	2.884***	2.965***
				(0.000)	(0.000)	(0.000)
OLS-BW						
r	-1.023***	-1.209***	-1.127***	-2.219***	-2.274***	-2.232***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
μ				12.357***	12.365***	12.129***
				(0.000)	(0.000)	(0.000)
observations	151538	163166	160310	151538	163166	160310

Notes: The table shows the results from a joint maximum likelihood estimation on the choices of the MPL and the OLS-BW based on the EUT-*crra* model without and with noise. The samples exclude violators of FOSD and SOSD. The OLS-BW choices are reweighed to make them comparable in the number to the MPl choices. P-values of the estimated coefficients are based on clustered standard errors and reported in parentheses. *** p<0.01; ** p<0.05; *p<0.1 significance level.

Table 3.11: Estimated Risk Preferences based on EUT-*crra* Model with Heterogenous Stochastic Noise

	(1)	(2)	(3)	(4)	(5)	(6)
model	EUT-*crra*	EUT-*crra*	EUT-*crra*	EUT-*crra*	EUT-*crra*	EUT-*crra*
experiment	*MPL*	*OLS-BW*	*MPL*	*OLS-BW*	*MPL*	*OLS-BW*
r	-0.556***	-2.246***	-0.556***	-2.232***	-0.518***	-2.273***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
μ_0	3.274***	6.326*	3.184***	5.734*	3.101***	9102.8***
	(0.000)	(0.067)	(0.000)	(0.079)	(0.000)	(0.000)
individual characteristics						
μ_{Age}	-0.00225	0.0786	-0.00169	0.0828	-0.00182	-0.0284
	(0.316)	(0.362)	(0.456)	(0.297)	(0.369)	(0.389)
μ_{Gender}	-0.0842	2.067	-0.0844	1.930	-0.0708	-1.390*
(1=female)	(0.168)	(0.281)	(0.162)	(0.299)	(0.203)	(0.036)
$\mu_{Literacy}$	-0.0402	5.273	-0.0285	5.274	-0.0295	0.769
(1=literate)	(0.480)	(0.141)	(0.618)	(0.120)	(0.580)	(0.260)
$\mu_{Livestock}$	-0.0169	-0.0469	-0.0174	-0.107	-0.0175	0.0944
	(0.411)	(0.917)	(0.406)	(0.805)	(0.322)	(0.567)
session-specific features						
μ_{time}			0.120**	0.956		
(1=afternoon session)			(0.021)	(0.561)		
enumerator dummies						
μ_{enu_1}					0.291**	-9087.5***
					(0.003)	(0.000)
μ_{enu_2}					0.0552	-9092.8***
					(0.588)	(0.000)
μ_{enu_3}					-0.0458	-9090.2***
					(0.639)	(0.000)
μ_{enu_4}					-0.570***	16651.0***
					(0.000)	(0.000)
μ_{enu_5}					0.0299	-9095.5***
					(0.740)	(0.000)
μ_{enu_6}					0.455***	10232.2***
					(0.000)	(0.000)
μ_{enu_7}					0.221*	-9093.1***
					(0.085)	(0.000)
μ_{enu_8}					-0.00903	-9091.4***
					(0.921)	(0.000)
μ_{enu_9}					0.149	-9069.4***
					(0.163)	(0.000)
observations	87278	856	87278	856	87278	856

Notes: The table shows the results from separate maximum likelihood estimations on the choices of the MPL and the OLS-BW based on the EUT-*crra* model, with stochastic noise explained by different observable characteristics. The estimation excludes participants who did not answer the survey questions after the experiments. P-values of the estimated coefficients are based on clustered standard errors and reported in parentheses. p<0.01; ** p<0.05; *** *p<0.1 significance level.

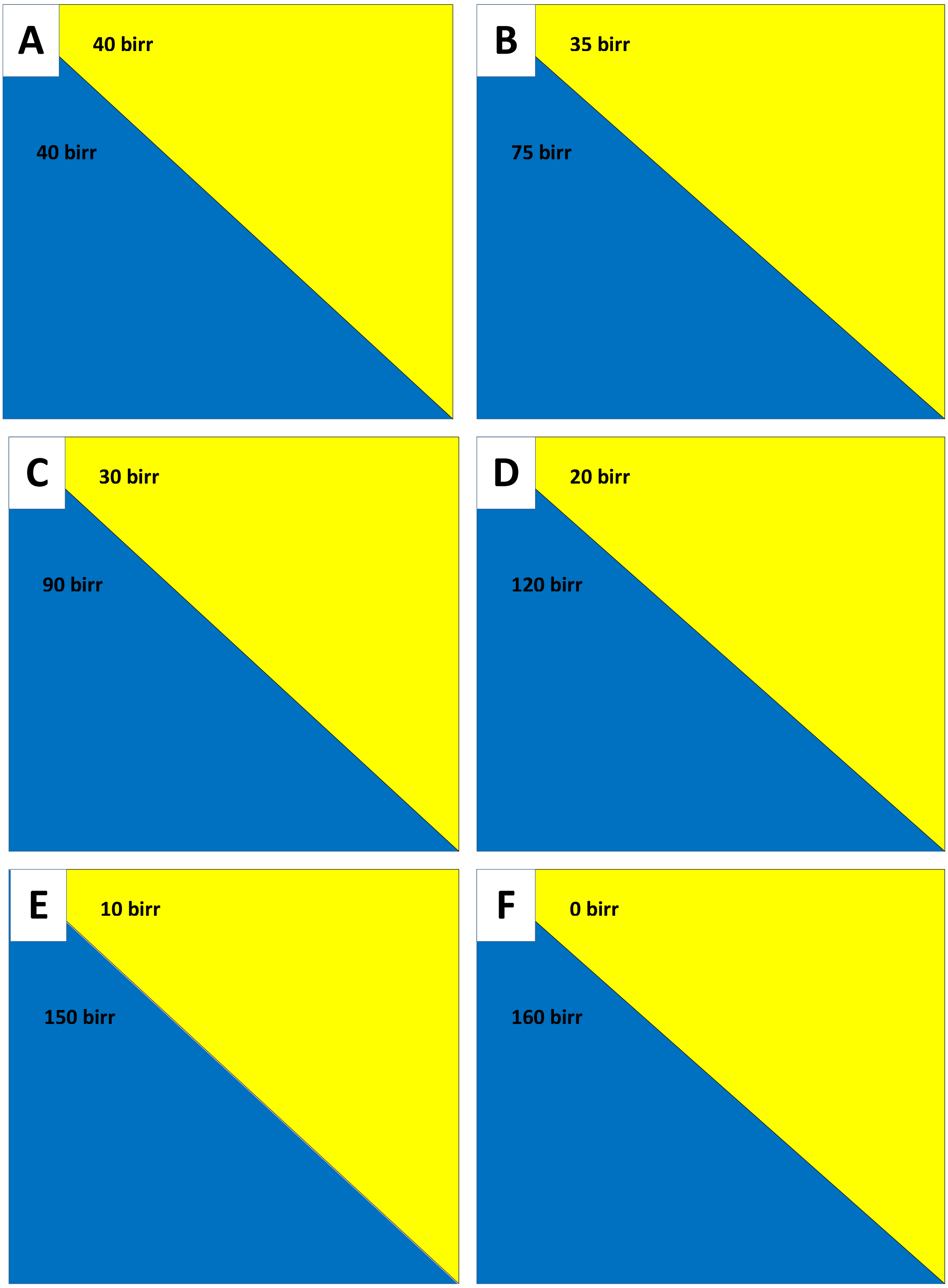
Figure 3.1: Price Lists in the MPL Experiment

price list 1					price list 2					price list 3				
Lottery	#	L	S	Sure	Lottery	#	L	S	Sure	Lottery	#	L	S	Sure
Win 160, p=0.5	1	O	O	0	Win 160, p=0.25	1	O	O	0	Win 160, p=0.75	1	O	O	0
	2	O	O	10		2	O	O	10		2	O	O	10
	3	O	O	20		3	O	O	20		3	O	O	20
	4	O	O	30		4	O	O	30		4	O	O	30
	5	O	O	40		5	O	O	40		5	O	O	40
	6	O	O	50		6	O	O	50		6	O	O	50
	7	O	O	60		7	O	O	60		7	O	O	60
	8	O	O	70		8	O	O	70		8	O	O	70
	9	O	O	80		9	O	O	80		9	O	O	80
	10	O	O	90		10	O	O	90		10	O	O	90
	11	O	O	100		11	O	O	100		11	O	O	100
	12	O	O	110		12	O	O	110		12	O	O	110
	13	O	O	120		13	O	O	120		13	O	O	120
	14	O	O	130		14	O	O	130		14	O	O	130
	15	O	O	140		15	O	O	140		15	O	O	140
	16	O	O	150		16	O	O	150		16	O	O	150
	17	O	O	160		17	O	O	160		17	O	O	160

price list 4					price list 5					price list 6				
Lottery	#	L	S	Sure	Lottery	#	L	S	Sure	Lottery	#	L	S	Sure
Lose 160, p=0.5	1	O	O	-160	Lose 160, p=0.25	1	O	O	-160	Lose 160, P=0.75	1	O	O	-160
	2	O	O	-150		2	O	O	-150		2	O	O	-150
	3	O	O	-140		3	O	O	-140		3	O	O	-140
	4	O	O	-130		4	O	O	-130		4	O	O	-130
	5	O	O	-120		5	O	O	-120		5	O	O	-120
	6	O	O	-110		6	O	O	-110		6	O	O	-110
	7	O	O	-100		7	O	O	-100		7	O	O	-100
	8	O	O	-90		8	O	O	-90		8	O	O	-90
	9	O	O	-80		9	O	O	-80		9	O	O	-80
	10	O	O	-70		10	O	O	-70		10	O	O	-70
	11	O	O	-60		11	O	O	-60		11	O	O	-60
	12	O	O	-50		12	O	O	-50		12	O	O	-50
	13	O	O	-40		13	O	O	-40		13	O	O	-40
	14	O	O	-30		14	O	O	-30		14	O	O	-30
	15	O	O	-20		15	O	O	-20		15	O	O	-20
	16	O	O	-10		16	O	O	-10		16	O	O	-10
	17	O	O	0		17	O	O	0		17	O	O	0

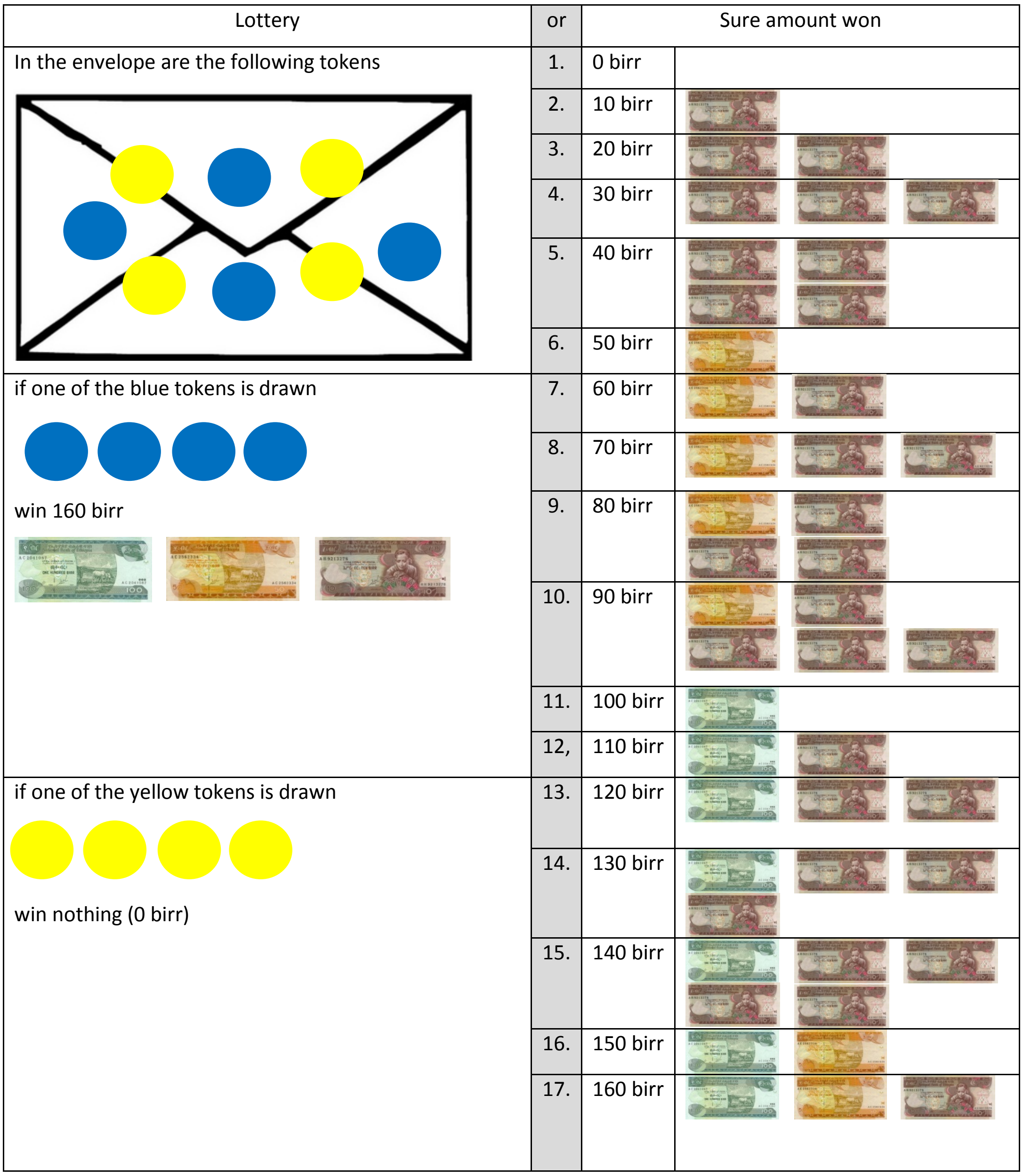
Notes: The figure shows the enumerators' form to note the respondents' choices. In each of the six price lists, respondents repeatedly choose between the lottery (L) and an increasingly sure amount (S) of winning or losing. The values in the lotteries and safe amounts are in ETB.

Figure 3.2: Illustration of the OLS-BW Experiment



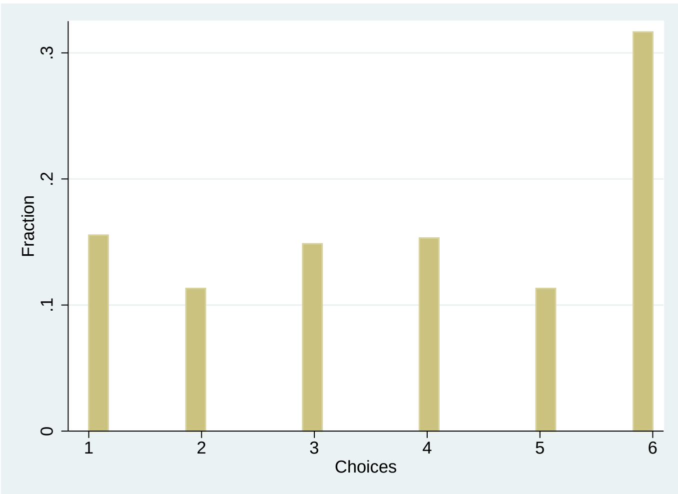
Notes: The illustration was shown at the tables, when participants made their decisions. Each participant decided for one of the six lotteries (A-F). The values in the lotteries are in ETB.

Figure 3.3: Illustration of the MPL Experiment: Price List with a Winning Chance of 50%



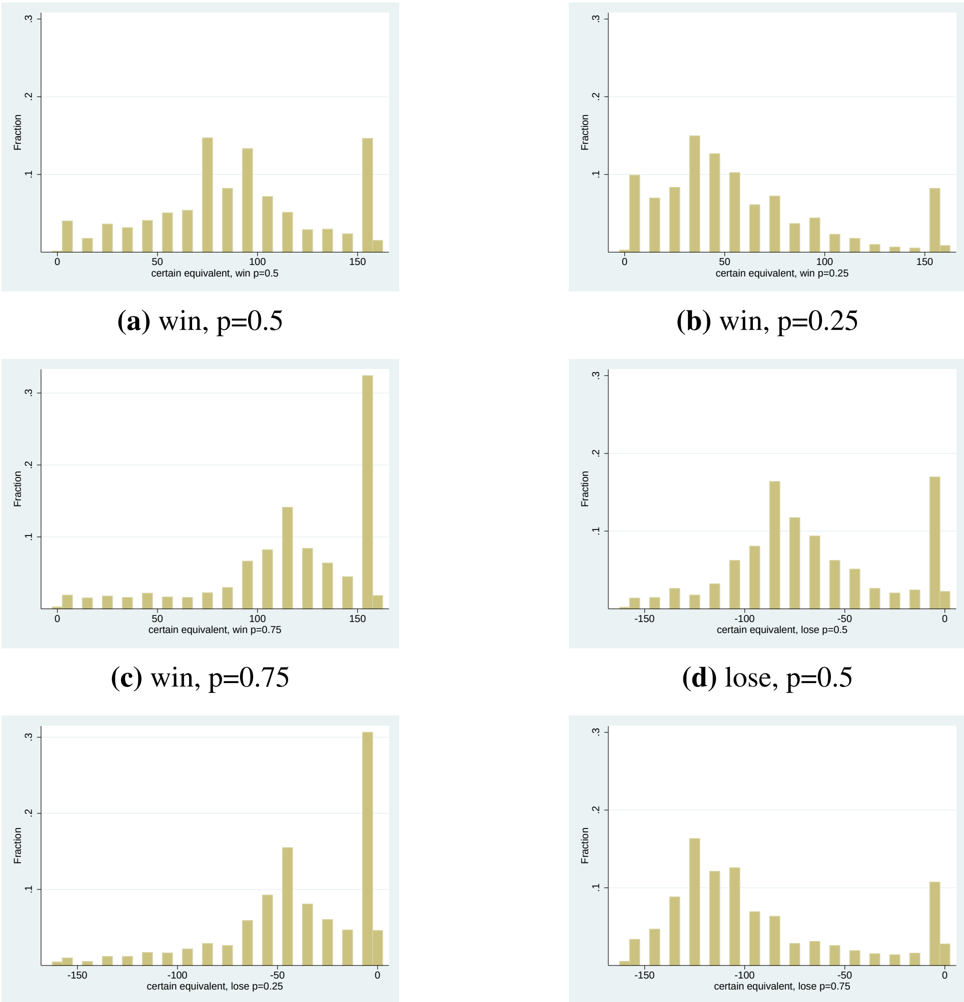
Notes: The illustration was shown at the tables, when particpants made their decisions. Each participant made 17 decisions between the lottery of winning 160 ETB with a chance of 50% and the increasingly sure amount. The values in the lotteries and safe amounts are in ETB.

Figure 3.4: Distribution of OLS-BW Choices



Notes: Respondents choosing the first lottery (A=1) with an expected payoff of 40 ETB are extremely risk averse, while respondents choosing the last lottery (F=6) with an expected payoff of 80 ETB are either risk neutral or risk loving.

Figure 3.5: Distribution of Certainty Equivalents in the MPL



(a) win, p=0.5

(b) win, p=0.25

(c) win, p=0.75

(d) lose, p=0.5

(e) lose, p=0.25

(f) lose, p=0.75

Note: CEs indicate the switching point in a price list. Respondents with the lowest CEs are extremely risk averse, while respondents with the highest CEs are extremely risk loving.

Chapter 4

Does Index Insurance Help Households Recover from Disaster? Evidence from IBLI Mongolia

Veronika Bertram-Hümmer
Kati Krähnert

4.1 Introduction

Index-based insurance was first introduced in the late 1990s in order to help agricultural households in developing countries cope with weather risks. Index insurance transfers covariate weather risks outside the local community and provides liquidity in the aftermath of a shock (Skees and Barnett 2006; Barnett and Mahul 2007; Barnett et al. 2008; Carter 2009). In the event of a weather disaster, insurance indemnity payments are expected to help households to recover from the damage caused by the weather shock. Despite the great enthusiasm with which index insurance is discussed in the policy community, there is little evidence of whether index insurance indeed provides its expected benefits (Miranda and Farrin 2012; World Bank Group 2013; Carter et al. 2014; Greatrex et al. 2015).

This paper investigates the impact of indemnity payments from index insurance on the asset recovery of households after a catastrophic weather disaster occurred. Our focus is on the Index-based Livestock Insurance (IBLI) in Mongolia. IBLI indemnity payments were triggered following the 2009/10 winter, when Mongolia was struck by a severe winter disaster (called *dzud* in Mongolian). Extremely cold temperatures, excessive snowfall, and an overall long duration of the winter caused the death of more than 10 million livestock, more than 23 percent of the national stock. In rural Mongolia, where the majority of the population depends upon herding activities for their livelihood, the 2009/10 dzud caused wide-spread poverty among herders (IFRC and MRCS 2010; Sternberg 2010). Our analysis builds on three waves of a household panel survey that we implemented in three western Mongolia provinces. We find a significant, positive and economically large effect of IBLI indemnity payments on households' post-disaster recovery in livestock. Thus, our study is among the first to document empirically a case in which index insurance works.

One challenge when evaluating the impact of a commercial product such as IBLI is that uptake is voluntary. Households voluntarily decide for or against purchasing IBLI, which is nationally advertised in the media and sold locally by insurance agents. Our empirical strategy accounts for selection into treatment with quasi-experimental methods. We employ the bias-corrected matching estimator to control for self-selection into purchasing IBLI before the disaster winter and receiving IBLI indemnity payments in 2010. Our analysis rests on the assumption that selection into treatment is independent of the outcome after controlling for observed covariates, with no unobservable factors remaining. Two factors help us to reduce a potential bias due to unobserved characteristics. First, we exploit the phasing-in of the IBLI program. When the dzud disaster of 2009/10 occurred, IBLI was still in its pilot stage and only available in one of the three survey provinces. We match treated households living in the province where IBLI was available in 2009 with control households living in the two other provinces where IBLI was not yet offered. This way, we increase the likelihood that treated and control households share similarities in unobserved characteristics. Moreover, the household survey questionnaire records detailed retrospective information on households' pre-shock herd size, livestock losses due to the winter disaster, as well as an unusually large number of household characteristics that we use to account for selection into treatment.

Our contribution to the small but growing empirical literature on the impact of index insurance on households is twofold. First, most existing studies focus on the *ex-ante* effects of insurance uptake on households' behavior (e.g., Giné and Yang 2009; Mobarak and Rosenzweig 2012; 2013; 2014; Karlan et al. 2014). This focus is due to the fact that index insurance was only first introduced in the late 1990s and in most empirical settings, a major weather shock that would have triggered indemnity payments – allowing for an analysis of the *ex-post* effects – has not yet occurred. A common conclusion in these empirical studies is that index insurance enables households to make larger and more profitable, but also riskier, investments. To the best of our knowledge, only one study by Janzen and Carter (2013) investigates the impact of indemnity payments from index insurance triggered by a drought. Janzen and Carter show that indemnity payments influence the shock coping strategies that households intend to use. We complement the study by Janzen and Carter by documenting that index insurance indemnity payments have an impact on observed asset dynamics of households after a weather disaster. Moreover, we use data on households' livestock holdings one, two, three and four years after the disaster and thus provide evidence on the persistence of the effect over time.

Second, our study is the first quantitative assessment of IBLI Mongolia, an index insurance program that is followed with great attention by policy stakeholders and insurance companies. Shortly after its implementation in pilot areas in 2006, IBLI had already created a sufficiently high demand (above 20 percent in some provinces) for it to be scaled up to the national level in 2012. It is the only index insurance scheme to date that has been transferred into an independent commercial insurance company. In contrast, most index insurance schemes that are implemented in developing countries are still in the pilot stage (Carter et al. 2014). So far, most of those schemes are either implemented on a small scale or heavily subsidized and thus not commercially viable. Almost all schemes struggle with low uptake (Binswanger-Mkhize 2012; Miranda and Farrin 2012; Carter et al. 2014).

Our results show a significant and positive effect of IBLI indemnity payments on households' recovery in livestock holdings in the first and second year following the 2009/10 winter disaster. This effect is less pronounced three and four years after the disaster, which indicates that in the middle term both insured and non-insured households were able to recover. Our results hold when defining the outcome variables in different ways (herd size and growth rates), when varying the numbers of matches per observation, when using different sets of covariates, when accounting for herd composition, and when using alternative propensity score estimators. Finding a consistently significant and economically large short-term effect of IBLI indemnity payments took us by surprise: Given that underinsurance is prevalent among IBLI customers (with households only purchasing insurance for a minor share of their livestock holdings) and given that our sample comprises a relatively small number of treated households, we did not expect to find strong results when beginning our analysis.

We further provide indicative evidence on the channels through which IBLI helped treated households to recover faster. This evidence is derived from analyzing households' shock coping strategies (again using the bias-corrected matching estimator) as well as complementary qualitative fieldwork in Mongolia. On the one hand, indemnity payments helped herders avoid selling and slaughtering animals, thus smoothing their productive asset base. On the other hand, IBLI appears to have relieved households from credit constraints, which may have been used to purchase new livestock.

The remainder of this paper is structured as follows: Section 4.2 provides a brief overview of existing empirical studies on index insurance. Section 4.3 provides contextual information on herding and weather risk in Mongolia and presents the design of IBLI. The household survey data are described in section 4.4, followed by an outline of the identification strategy in section 4.5. Section 4.6 presents estimation results and robustness tests as well as a discussion on the potential channels. Section 4.7 concludes.

4.2 Literature on the Impacts of Index Insurance

Given that index insurance is a relatively new product, the empirical literature on the impact of index-based insurance on households is still small. Existing research mostly focuses on the *ex-ante* effects of insurance uptake on households' risk management behavior. A common finding is that index insurance enables agricultural households to make riskier investments that generate higher returns. For instance, Karlan et al. (2014) find that farmers in northern Ghana who purchased index insurance cultivate more land, increase their efforts in preparing the land, and spend more on fertilizers. In a study on index insurance offered to cotton cooperatives in Mali, Elabed and Carter (2014) show that insured households cultivate a larger area of land with cotton and invest more in seeds. Hill and Viczeisza (2012) carry out an experimental game with Ethiopian farmers who are offered a stylized index insurance. Results indicate that index insurance increases the likelihood that farmers purchase fertilizer. In a series of papers, Mobarak and Rosenzweig (2012; 2013; 2014) show that even when informal risk-sharing networks are present, index insurance entices farmers in India to take greater risks. Insured farmers are more likely to plant a portfolio of rice varieties that are less drought-resistant, but generate higher yields. In a study of rural households in Andhra Pradesh, Cole et al. (2013) find that insured farmers allocate more agricultural inputs to the production of crops that generate higher expected returns, but are more sensitive to deficient rainfall. All of the studies are based on randomized experiments to identify the effects of index insurance. Moreover, all of the studies either focus on small index insurance schemes that are heavily subsidized or use stylized index insurance in the context of experimental games.

Reviews on index insurance conclude that there is a prevalent knowledge gap on the *ex-post* impacts of index insurance (Miranda and Farrin 2012; Carter et al. 2014). We are aware of only one existing study, by Janzen and Carter (2013), that explores the effectiveness of insurance indemnity payments in the aftermath of a weather shock. Janzen and Carter focus on an index-based livestock insurance pilot scheme in the Marsabit district of northern Kenya. The authors analyze the impact of indemnity payments triggered by a severe drought in 2011, using randomly distributed information and price discount treatments to identify selection into purchasing insurance. The majority of insured households purchased index insurance under subsidized rates. The outcome of interest is households' anticipated use of shock coping strategies in the next quarter of the year. This information is recorded during the drought, but before the insurance indemnity payments are transferred to households. Janzen and Carter find that insurance indemnity payments significantly affect the choice of coping strategies that households expect to use: Insured households are less likely to anticipate selling livestock and reducing meals compared to non-insured households. Moreover, the authors distinguish between households below and above a critical asset threshold. Insured households below the asset threshold are less likely to expect reducing the number of meals, while insured households above the threshold are less likely to expect selling livestock. Our study builds on the findings of Janzen and Carter (2013) by investigating the *observed* post-shock asset dynamics of Mongolian herders after receiving insurance indemnity payments.

4.3 Empirical Context

4.3.1 Herding and Weather Risk in Mongolia

Herding is the main economic activity in rural Mongolia. It is the single most important employment sector, accommodating about 35 percent of the workforce in 2012. Most households living outside the capital of Ulaanbaatar own livestock, with approximately 146,000 households consisting of about 545,000 individuals (about 19 percent of the population) deriving their livelihood from herding (NSO 2013). Extensive production techniques are commonly used in herding, with animals being grazed on open rangelands that are in state property. The majority of herders is nomadic or semi-nomadic and moves their herds between two and 25 times per year, usually returning to the same campsites every year. Herders typically own a mix of five species that are adapted to the extreme continental climate of Mongolia: sheep, goats, horses, cattle and camel. Sheep provide most of the meat for households' subsistence needs as mutton is the preferred type of meat. Cattle primarily provide milk that is consumed and used for other dairy products. Cashmere wool derived from goats is an important source of cash income. Horses and camels are mainly used for tending smaller livestock and for transportation; they are also considered a prestigious form of storing wealth. All animal species are sold (alive, slaughtered, as well as their skins and hides) if need arises. In 2014, herding households owned on average 244 animals. The national herd was estimated at 51.9 million animals (Mongolian Statistical Information Service 2015).

Mongolian herders face a number of risks and hazards that pose a constant threat to their livelihood. The most prevalent risk is extreme winter conditions, called dzud in Mongolian, which cause mass livestock losses. Dzuds are triggered by various and rather different climatic conditions, often by a combination of several (Batima 2006, p. 57; Murphy 2011, p. 32-33). Among them are too little precipitation (either in the preceding summer or during the winter) that limits vegetation growth; excessive snowfall that prevents animals from grazing; extremely cold temperatures that sharply raise the calories intake required by animals; and fluctuations in winter temperatures above and below freezing that lead the snow to melt and then ice over, thus making it difficult for animals to reach the grass. Dzuds are reinforced by local geographic features, such as the ecological zone, altitude, and the location on a slope, thus making it difficult to predict when and where dzuds will occur. Since 1990, dzuds occurred in the winter of 1999/00, 2000/01, 2001/02, and 2009/10.

Our focus in this paper is on the dzud of 2009/10, which caused the highest livestock mortality in a single winter ever recorded in Mongolia. This point is illustrated in figure 4.1, which presents livestock dynamics in the three survey provinces of Uvs, Zavkhan and Govi-Altai of western Mongolia between 1970 and 2014, calculated from annual livestock census data. Two aspects are noteworthy. First, the 2009/10 dzud pushed livestock numbers back to 1970 levels. Thus, a single catastrophic winter was enough to eliminate decades of livestock development. Second, there was a rapid and steady recovery in livestock starting in 2010. By 2014, livestock levels had almost reached 2009 levels.

The socio-economic consequences of the 2009/10 dzud were numerous. Some 40 percent of all herding households lost more than half of their herd (UNDP and NEMA 2010, p. 41). Many households were pushed below the herd size of 100 animals, which is considered the minimum necessary for sustaining a pastoralist livelihood in the long term. A sizeable number of impoverished herders moved as distress migrants to urban centers in search of employment (Sternberg 2010). The food security of severely affected households was threatened (IFRC and MRCS 2010). In turn, malnutrition experienced during the dzud months had lasting impacts on the human capital of children from herding households: Children who lived in severely affected regions were significantly shorter three years after the dzud compared to same-aged children living in less affected regions (Groppo and Schindler 2014).

While both the government and international agencies provided emergency aid to dzud-affected regions in 2010, this support has been ad hoc and did not reach all of the affected herders in the sparsely populated countryside (UNDP and NEMA 2010). When the 2009/10 dzud hit Mongolia, formal financial markets were only starting to develop in rural areas (Goodland et al. 2009). Thus, herders mainly drew on informal strategies to manage risk and cope with the consequences of dzuds (Fernández-Giménez et al. 2015). Increasing the herd size was the most important informal risk management strategy to prepare for harsh dzud winters (Goodland et al. 2009). Other common strategies applied in the midst of dzud winters include conducting additional nomadic movements and providing animals with supplementary fodder (mostly hay) (Murphy 2011; Fernández-Giménez et al. 2015). Yet, given the severity and covariate nature of dzuds, the effectiveness of informal risk management mechanisms is limited. As a consequence, “high levels of livestock mortality are often unavoidable even for the most experienced herders” (Mahul and Skees 2007, p. 10). After dzud, restocking is the most important goal for herders (Goodland et al. 2009).

4.3.2 Index-based Livestock Insurance Mongolia (IBLI)

After three consecutive dzud winters between 1999 and 2002, there was consensus among policymakers in Mongolia that herders needed effective and sustainable insurance against livestock losses caused by dzuds (Mahul and Skees 2007). The Government of Mongolia requested assistance from the World Bank to create a livestock insurance scheme that suited the characteristics of the Mongolian herding sector. The World Bank proposed IBLI that would be offered to herders as a commercial product by private insurance companies (Skees and Enkh-Amgalan 2002).

In 2006 IBLI was introduced as a pilot project in three – Bayankhongor, Khentii, and Uvs – out of the 21 provinces of Mongolia (see the map in figure 4.2). During the pilot phase, the Project Implementation Unit (PIU) of IBLI played a key role in organizing and marketing the insurance as well as in supporting insurance agents selling actual IBLI policies. Insurance uptake rates increased steadily over the pilot phase and were as high as 27 percent in some provinces (Hartell 2011, p. 27). After the encouraging performance during the pilot phase, IBLI was stepwise scaled up to the national level until 2012. IBLI has since been offered in every district of Mongolia. In 2014, all PIU activities were handed over to the participating private insurance companies and the provincial PIU offices were closed. At the same time, a reinsurance company with public-private ownership was established to serve as reinsurer for the private sector insurance companies selling IBLI (Law of Mongolia 2014).

The key objective of IBLI is "to reduce herders' livelihood vulnerability caused by dzud or natural disasters" (PIU 2012, p. 12). In more general terms, IBLI aims at improving herders' welfare by increasing herd size, assets, savings, income, and contribution to income from herding (Hartell 2011, p. 11). Under IBLI, indemnity payments are made to insured herders when the district-level mortality rate of a given livestock species exceeds the threshold of 6 percent.[1] Herders receive indemnity payments irrespective of the actual losses they personally experienced (which would be very costly to verify, given the

[1] Two data sources are used to calculate livestock mortality rates: First, the total number of adult animals is obtained from the yearly livestock census, which has been implemented since the 1920s. Every year in mid-December, the National Statistical Office of Mongolia (NSO) gathers data on the national stock. This exercise is carried out collaboratively by enumerators and local officials, who maintain detailed records of herders and their livestock in their administrative division. Second, in June of each year, the NSO conducts a livestock survey to establish the losses of adult animals due to dzud during winter and spring. From the

vast distances in Mongolia's countryside). Moderate losses that fall below the 6 percent threshold are not covered by IBLI and remain the responsibility of the herders alone. This feature ensures that herders continue implementing informal risk management strategies, thus discouraging moral hazard (Hartell 2011). More severe shocks resulting in livestock loss rates above the 6 percent threshold trigger IBLI indemnity payments, which are paid from a joint-liability pool of the participating private insurance companies. One challenge inherent in the nature of dzuds is that extremely severe shocks occur every now and then. Such catastrophic covariate events are likely to trigger indemnity payments in many districts, which may render local insurance companies bankrupt. In order to make IBLI sustainable in the long term, the government and the global reinsurance market cover indemnity payments if mortality rates exceed 30 percent.

For herders, participation in IBLI is voluntary. There are several features inherent in IBLI that allow herders to customize IBLI to their needs. Premium rates differ across districts and species, reflecting the local livestock mortality risk.[2] Herders can choose purchasing IBLI for any of the five common livestock species; for instance, herders can insure only horses, but not their cattle. Herders also decide on the insurance value for each species, which can be between 1 and 100 percent of the animals' market value. Incomplete insurance is prevalent, with herders insuring on average only 30 percent of the value of their herd (PIU 2012).

Local insurance agents representing the six Mongolian insurance companies participating in the IBLI program sell IBLI policies to herders. IBLI policies are sold between April and June in a given year and cover livestock losses occurring between December and June of the following year. Given that the sales period ends in mid-summer, neither herders nor insurance companies can predict conditions in the next winter, which prevents adverse selection. Indemnity payments are made to insured herders from August the following year onwards. The processing of insurance premiums and indemnity payments is done via local banks. Some banks also offer loans with discounted interest rates to insured herders (PIU 2012).

two data sources, the district-level mortality rates of adult animals are then calculated separately for each species.

[2]For example, a herder in Tunel district of Khuvsgul aimag wishes to insure his 18 horses under IBLI: The market value for a horse in Khuvsgul aimag is 225 US$ and the IBLI premium rate for a horse in Tunel district is 1.69 percent. The total livestock value is 18*225 US$=4,045 US$. If the herder decides to insure his horses at 100 percent market value, the insurance premium would be 1.69 percent*4,045 US$=68 US$. Imagine the district-level mortality of horses in Tunel district is 15 percent in the following winter. Then the indemnity rate would be 15 percent-6 percent=9 percent. The herder would receive an indemnity payment of 9 percent*4,045 US$=364 US$ (example taken from PIU 2012, p. 48).

Our focus in this paper is to explore if IBLI helped insured households cope with the catastrophic dzud of 2009/10, which occurred while IBLI was still in its pilot phase. In 2009, IBLI was available to herders in four provinces, including the province of Uvs in which the household survey was implemented. In Uvs province, 1,835 herders purchased IBLI in 2009, representing an uptake rate of 19.5 percent.[3] On average, herders in Uvs province insured 102 heads of livestock and paid an insurance premium of 28,000 Mongolian Tugrik (MNT) (about 19 US$).[4] The losses in the winter 2009/10 triggered indemnity payments to 95.4 percent of insured herders in Uvs province, who received on average 416,000 MNT (about 312 US$) as indemnity payment (PIU 2012). The large overall amount of indemnity payments that were triggered by the 2009/10 dzud posed a challenge for the sustainability of the IBLI scheme. Additional support from the World Bank and other donors was necessary to stem the extremely high amounts of indemnity payments after the 2009/10 dzud disaster.[5]

4.4 Data

Our analysis builds on three waves of the *Coping with Shocks in Mongolia Household Panel Survey* that is implemented by the authors in collaboration with the National Statistical Office of Mongolia (NSO). The survey data are collected in the three neighboring provinces (aimags) of Uvs, Zavkhan, and Govi-Altai in western Mongolia (see figure 4.2; the survey provinces are bold-rimmed) and cover 49 out of 61 districts (soums) in these three provinces.[6] Compared to other regions of Mongolia, the three survey provinces resemble each other in terms of socio-economic characteristics of the population, economic activities, and the large distance to the capital Ulaanbataar, which

[3]The household survey data comprise 261 herding households that lived in Uvs in 2009, of which 59 households purchased IBLI. Thus, the IBLI update rate in our sample of 22.6 percent corresponds closely to the actual uptake rate in the population.

[4]The exchange rate was about 1 US$=1,470 MNT during the time of the winter disaster in 2009/10.

[5]As a result of the 2009/10 dzud, the insurance product was slightly revised. Until 2010, herders could choose between the Disaster Risk Product (DRP), which covered only catastrophic losses above a threshold of 30 percent, and the Base Insurance Product (BIP), which covered all losses above the triggering threshold of 6 percent (Mahul and Skees 2007; Miranda and Farrin 2012). After the 2009/10 dzud, the DRP was abolished both due to low uptake and unsustainable reinsurance arrangements. The BIP was transformed into the so-called Livestock Risk Insurance (LRI), which covers losses above the 6 percent and the 30 percent threshold within one product (DeAngelis 2013). In the analyses presented in the following, we do not distinguish between herders buying DRP or BIP in 2009.

[6]A province is the top level of Mongolia's administrative structure. Each province is subdivided into several districts (*soums*), which are further subdivided in sub-districts (*bags*). There are 21 provinces, 329 districts, and 1,720 sub-districts in Mongolia. As of 2011, districts in western Mongolia have an average population of 3,154 persons and a size of 4,811 square kilometers.

is a good proxy for access to urban markets and economic opportunities.

The survey sampling is based on the Population and Housing Census of 2010 and uses a multi-stage design, which ensures that the sample is representative of the population in western Mongolia.[7] More specifically, statistically significant results are achieved for each of the three survey provinces, for urban areas in each province, and for rural areas in each province. The sampling design was not influenced in any way by the purpose of conducting an impact study of IBLI. Rather, the aim of the survey was to gather representative data of the population in the region, which underlines the robustness of our results. The first, second, and third panel waves were collected between June 2012 and May 2013, between June 2013 and May 2014, and between June 2014 and May 2015, respectively. For the sake of readability, we identify variables collected in the first, second, and third panel waves as 2012, 2013, and 2014. Household interviews were conducted continuously throughout the year, with one-twelfth of the sample households interviewed every month. The data are representative across seasons.

The sample comprises 1,094 households owning livestock in 2009. All sample households that purchased IBLI in 2009 reported experiencing the dzud and all of them were still herders when the household survey was implemented. To ensure that the group of non-insured control households is comparable to the insured treated households, we exclude 125 control households that reported not being exposed to the dzud and 118 control households that lost their entire herd during the 2009/10 dzud and dropped out of the herding economy in the aftermath of the dzud. Moreover, we exclude all 206 non-insured households living in the province of Uvs. Lastly, we exclude 3 control households with missing values in key covariates. This leaves us with a sample of 642 herding households.

The survey questionnaire collects detailed information on the demographics of each household member, household consumption expenditures, income, assets, subjective well-being, social networks, livestock holdings, strategies in herding as well as infrastructure and environmental conditions in the district. One questionnaire module focuses specifically on households' exposure to the 2009/10 dzud and post-dzud coping strategies

[7] In the first sampling step, the three provinces were subdivided into nine mutually exclusive strata of province centers, district centers, and rural areas. In the second step, Primary Sampling Units (PSU) were randomly drawn from each stratum, resulting in a total number of 221 PSUs. In a third sampling step, inside each PSU eight households were randomly selected. Unfortunately, the matching estimator as well as other propensity score estimators that we draw upon do not allow to account for survey design effects and the clustering of standard errors (cf. Guo and Fraser 2014, p. 243). To assess the impact of design effects, we estimate two versions of our main specification with OLS: one simple version and one version that accounts for design effects and clustering of standard errors. Results (available upon request) are almost identical, which makes us confident that this issue is less of a concern.

applied.

Another questionnaire module records detailed information on IBLI from all herding households. This module asks for retrospective information on the purchase of IBLI in the past. More specifically, information is available on whether the household purchased IBLI in 2009, the amount of indemnity payments received in 2010, how indemnity payments were used, and whether the household found the indemnity payments helpful.[8] It is important to note that the treatment variable in our analysis – whether a household had purchased IBLI in 2009 – is recorded twice in two separate modules of the questionnaire (in the module on IBLI and in the module on the 2009/10 dzud). All but two treated households gave consistent answers on their insurance status in 2009 in these two modules, which underlines the reliability of the treatment variable. Our sample comprises 59 treated households and 583 control households.

The key variable in our analysis is households' livestock holdings at five points in time: in 2009 (before the shock) as well as in 2011, 2012, 2013, and 2014.[9] In each of the three panel waves (2012, 2013, 2014), households are asked about their current livestock holdings at the time of the survey interview. The livestock holdings in 2011 are obtained from the first panel wave, when households were asked to also report their livestock holdings 12 months earlier. The livestock holdings in 2009 (and also livestock losses experienced during the dzud, which is included as covariate in our multivariate analysis) are asked retrospectively from households.

Two empirical observations make us confident that the retrospectively recorded information on past livestock holdings is reliable. On the one hand, households are asked about their livestock holdings in 2009 and dzud-related livestock losses twice, in the first panel wave and again in the third panel wave. The coefficient of correlation for livestock holdings in 2009 recorded in the first and third waves is 0.79; the coefficient of correlation for livestock losses is 0.83. Our preferred measure is retrospective information recorded in the first panel wave, given that the recall period is shorter. As a robustness test, we carry out all estimates with retrospective information recorded in the third wave and obtain very similar results (as will be discussed in section 4.6.3). This test assures us that

[8] Unfortunately, information is not available on the number and species of livestock insured in 2009 and on the amount of insurance premium paid in 2009.

[9] Household interviews for each panel wave were carried out over a period of 12 months, starting in June of each year. To ensure comparability in herd size across households interviewed before and after the birthing season, we exclude newborns born between January and May.

the retrospectively recorded variables are of good quality.

On the other hand, anthropological studies on Mongolia stress the importance of livestock holdings for the social standing of households. According to Murphy, "being wealthy in livestock, to be *myangat* or *bayan*, carries additional symbolic power beyond the economic value of the herd. It is more than being *ix maltai* (having many stock)" (Murphy 2011, p. 132). For instance, there are specific terms in the Mongolian language to classify herders with different livestock holdings ($<$100 heads; 100-200; 200-500; 500-1,000; and $>$1,000 heads) (ibid.). Therefore, it is not surprising that our survey enumerators did not observe difficulties among respondents to recall the size of their herd in the past. In fact, the questionnaire also asks for livestock holdings even further back in time (for instance, in 1990, 1999, and 2002), which respondents also reported without difficulty.

Complementary to the household survey data, we employ aggregated data from the Mongolian livestock census, which is implemented annually in mid-December by the NSO. More specifically, we use data on adult livestock mortality in 2010 at the level of the sub-district.

4.5 Identification Strategy

The aim of our analysis is to explore the causal effects of receiving IBLI indemnity payments in 2010 on households' post-disaster livestock recovery. We need to tackle the problem of missing data on the counterfactual: What would livestock recovery of insured households have been in the absence of indemnity payments? The IBLI scheme was implemented without randomized assignment rules at the household level. Thus, our empirical strategy needs to account for selection into treatment. All 59 households in our sample that purchased IBLI in 2009 also received indemnity payments in autumn 2010. Therefore, the treatment is an indicator variable taking the value one if a household purchased IBLI in 2009. Our identification strategy exploits the phasing-in of IBLI, with IBLI initially available in only one of the three survey provinces in 2009. The treated sample households all live in the province of Uvs, whereas control households live in the provinces of Zavkhan and Govi-Altai, where IBLI was introduced only after the dzud winter. We then employ matching methods to simulate a counterfactual.

More specifically, we employ the bias-corrected matching estimator (Abadie and Imbens 2002; 2006; 2011) to account for selection into purchasing IBLI in 2009 based on observable characteristics. The matching estimator uses a vector norm to impute a counterfactual outcome for each sample household (both the treated and non-treated). The vector norm calculates distances on the covariates between a treated household and each of its potential control households and vice versa (Guo and Fraser 2014, p. 212). Then the outcomes of households of the other treatment status that exhibit the shortest distance in covariates – the nearest-neighbors – are imputed as counterfactual outcomes to each household.

Formally, for each household h we know one of the potential outcomes, namely $Y_h(0)$ for those who purchased IBLI in 2009 and $Y_h(1)$ for those who did not. If a household purchased IBLI ($T_h = 1$), we use its observed outcome Y_h as its estimated outcome under treatment, such that $Y_h = Y_h(1) = \tilde{Y}_h(1)$. If a household did not purchase IBLI (and therefore has not received indemnity payments) ($T_h = 0$), we use the observed Y_h as our estimate under non-treatment, with $Y_h = Y_h(0) = \tilde{Y}_h(0)$. To derive the estimates of the counterfactual outcomes, the matching estimator imputes the non-observed outcomes for each household h using the average of the outcomes of the M-closest households in the opposite treatment group J. The estimated outcomes in the treatment $\tilde{Y}_h(1)$ and the non-treatment $\tilde{Y}_h(0)$ case are then estimated as simple matching estimator (Abadie and Imbens 2002):

$$\tilde{Y}_h(1) = \begin{cases} \frac{1}{M}\sum_{j\in J_M(h)} Y_j & \text{if} \quad T_h = 0 \\ Y_h & \text{if} \quad T_h = 1 \end{cases} \tag{4.1}$$

$$\tilde{Y}_h(0) = \begin{cases} Y_h & \text{if} \quad T_h = 0 \\ \frac{1}{M}\sum_{j\in J_M(h)} Y_j & \text{if} \quad T_h = 1 \end{cases} \tag{4.2}$$

Our set of covariates includes several continuous covariates (discussed below), which would result in inaccurate matching and lead to biased estimates (see Abadie and Imbens 2006 for a discussion). Therefore, we additionally include a bias-correction term when imputing the counterfactual outcomes. The bias correction is based on a linear regression of Y_j on the covariates X_j of the M-closest control observations j. It adjusts the counterfactual estimates for the differences in the covariate values for each observation X_h with its matched observations from the opposite treatment group X_j. Matching is done with replacement: each household is matched to several households of the opposite treatment status. This feature maximizes the number of matches used in the analysis and is thus well-tailored to the small number of treated households in our empirical setting. We match each household with four households (*M*=4) of the other treatment status, following recommendations by Abadie et al. (2004).[10] We estimate the impact of indemnity payments as sample average treatment effect (ATE).[11]

The matching estimator builds on the following three key assumptions: First, it assumes that the selection into treatment is independent of the outcome after controlling for observed covariates (i.e., the ignorable treatment assignment assumption holds). If there are remaining unobservables that are correlated with both treatment and outcome, the matching estimator yields biased estimates. As we cannot exploit any household-level randomization in the implementation of IBLI, we are unable to control explicitly for selection based on unobservables. Yet, exploiting the phasing-in of IBLI allows us to partly account for selection based on unobservables. Non-treated households in Uvs province (where IBLI was available in 2009) are excluded from the sample, while the control group consists only of households living in provinces where IBLI was not yet available in 2009. This way, we increase the likelihood that treated and control households share similarities in unobserved characteristics. Moreover, our household survey data include an unusually large number of pre-shock and time-invariant variables that we include as covariates in our estimations. Given that at least some of the unobserved characteristics are likely to be correlated with observed characteristics, we hope to keep this bias to a minimum.

The second assumption requires a sufficient overlap in the distribution of covariates between treated and non-treated sample households. When the overlap condition is fulfilled, each household has a positive probability of receiving each treatment level. We use propensity score methods to investigate the overlap for three set of covariates (explained later on in this section). Figure 4.4 shows that the overlap in the propensity scores is sufficient for all three sets of covariates. It is quite large for the minimal set and the core set of covariates and still acceptable in the maximal set of covariates.

[10]The estimations were carried out using the *teffects nnmatch* command in Stata. The Breusch-Pagan and Cook-Weisberg tests for heteroscedasticity reject the null hypothesis of constant variance for most covariates. Therefore, we specified a robust variance estimator allowing for heteroscedasticity. As variance matrix, we use the inverse of the variance-covariance matrix, with the Mahalanobis metric to calculate distances.

[11]The ATE calculates the average treatment effect of indemnity payments on all herding households in our sample. This choice reflects the fact that IBLI aims at targeting all herding households and uptake rates were relatively high.

Third, the estimated coefficient of the treatment variable only reflects the true effect if the stable unit treatment value assumption (SUTVA) holds. The potential outcome of each household needs to be independent of the potential outcomes of all other households. This assumption would be violated if indemnity payments were distributed within herders' networks, for instance within the family or across households that share a campsite. Such patterns could lead to general equilibrium effects that may affect the post-shock recovery of all households – also of the non-insured. Our study minimizes the possibility of a violation of SUTVA by exploiting the phasing-in of the IBLI scheme across provinces.

We employ two types of outcome measures, with each defined for several years. First, we use post-disaster herd size in the years 2011, 2012, 2013, and 2014. Second, we use livestock growth rates, which explicitly take into account the pre-shock herd size. More specifically, we construct the cumulative annual livestock growth rates[12] in the periods 2009-2011, 2009-2012, 2009-2013, and 2009-2014.

We employ three sets of covariates when estimating the effect of IBLI indemnity payments on post-shock livestock recovery. This way, we account for the fact that the choice of covariates can influence the estimated results when using methods that correct for self-selection into treatment (Guo and Fraser 2014). The full list of control variables and summary statistics are displayed in table 4.1.

The first set of covariates is informed by a probit estimation of the determinants of purchasing IBLI in 2009 (see table 4.2). The sample consists of all herding households living in the province of Uvs in 2009. Of the various controls included, only households' livestock holdings in 2009 and the local ecological zone in the district have a statistically significant effect on the purchase decision. This result is in line with our expectations and suggests that it is primarily exposure to risk factors that influence the uptake of IBLI: Households with large herds have fewer alternative income sources and are, thus, more vulnerable to livestock losses. The ecological zone is a good proxy for the long-term risk of experiencing a dzud. In column 2, we additionally include district fixed effects, which controls most comprehensively for any differences related to the supply of IBLI and the risk of facing dzud across districts. Livestock holdings in 2009 remain the only significant predictor for purchasing IBLI in 2009. Based on these findings, the first set of covariates

[12] The cumulated growth rate is a geometric progression ratio. For instance, the 2009-2012 growth rate is defined as $(\frac{herd_{2012}}{herd_{2009}})^{\frac{1}{3}} - 1$. It represents the average annual growth rate, assuming that livestock dynamics were constant across years. This definition is commonly applied in the financial economics literature, e.g. when calculating expected average returns of investments over time (Feibel 2002).

(referred to as *minimal set of covariates* below) comprises households' livestock holdings in 2009 as well as characteristics of the local ecological zone. Moreover, we include in the minimal set of covariates the number of livestock lost during the 2009/10 dzud as self-reported by households.

The second set of covariates (referred to as *core set of covariates* below) additionally consists of variables that are also identified by the existing theoretical and empirical literature on index insurance to explain selection into purchasing insurance. For instance, empirical studies by Cole et al. (2014) and Giné et al. (2008) find that wealth and liquidity are important determinants of index insurance uptake in India and Kenya, respectively. We account for these factors by controlling for pre-shock herd size, which is both the most adequate measure of wealth among herders and a suitable proxy for cash income from agriculture.[13] Moreover, we utilize households' subjective relative well-being in 2009. More specifically, households are asked to rank their socio-economic position just before the 2009/10 dzud started relative to other households in their district on a scale from 0 to 10. Compared to other studies based on observational data, the availability of retrospective information on wealth is a unique feature of the data at hand.

Attitude toward risk is also identified as a key factor influencing the decision to purchase index insurance (Clarke 2011; Cole et al. 2014). Hence, we account for the risk attitude of the head of household. The survey questionnaire asks respondents to indicate their general willingness to take risk on a 0-10 scale. This way of eliciting risk preferences works reasonably well when used as a control variable in empirical studies with samples from a developing country (Hardeweg et al. 2013; Nielsen et al. 2013). Furthermore, Gaurav et al. (2011) show that financial literacy matters for the decision to insure, given that index insurance is a complex financial product. We proxy financial literacy with the education and age of the head of the household.

Lastly, shock intensity and shock coping opportunities are important predictors of recovery (Janzen and Carter 2013). In addition to the household-level shock measure, we use livestock mortality in 2010 at the sub-district level (calculated from livestock census data) to proxy for the covariate nature of the dzud. An indicator variable taking the value one if households live in rural areas (as opposed to district centers and provincial centers) proxies income opportunities outside the herding economy. At the district level, the availability of cellphone networks and the number of public transportation options to

[13] In 2012, the coefficient of correlation between herd size and household cash income from herding was 0.68.

the provincial capital are employed to capture differences in economic access and hence in opportunities for post-dzud recovery across districts.[14]

The third set of covariates (referred to as *maximal set of covariates* below) additionally includes three controls that proxy households' ability to cope with dzud. These include the percent of female breeding stock out of total herd size and the number of economically active household members. During a harsh winter, the demand for labor in herding activities increases. For instance, livestock needs to be tended more closely and moving the herd to areas less affected by the dzud is highly labor intensive. In addition, we employ an indicator variable taking the value one if a household knows the local sub-district governor very well. A household that maintains close ties to the local governor might have a larger social network within the community and thus more options at hand to receive support (both informal and formal) following the dzud. While these three variables capture important differences across households, they entail one limitation: All three variables refer to 2012, the time of the first panel interview, which makes them potentially endogenous to the household-level shock exposure. Therefore, their inclusion as covariates may be disputable.

One potential concern in our analysis is that non-treated herders may have been exposed to a higher dzud intensity in the 2009/10 winter than the treated households (Hartell 2011). If treated households faced less severe dzud conditions in the winter 2009/10 and hence had more favorable circumstances for post-dzud recovery than non-treated households, we would falsely attribute their faster post-disaster recovery to IBLI indemnity payments. We undertake three measures to minimize this concern. First, recall that our estimations include controls for the intensity of the 2009/10 dzud at the household level and the sub-district level. Second, we explore whether treated and control households differ in the number of livestock lost during the dzud after controlling for observable household and district characteristics. Results of an OLS estimation of the determinants of livestock losses due to the dzud are displayed in table 4.3. The estimated coefficient of purchasing IBLI in 2009 is not statistically significant. Hence, there is no indication that treated and control herders experienced a statistically different shock intensity. Third, we explore the spatial variation in dzud intensity across sub-districts. Figure 4.3 shows the livestock mortality in 2010 in sub-districts included in the household survey. Clearly, the variation in dzud intensity within each province is very large: In each province, there are sub-districts experiencing low livestock mortality (below 16 percent) and high livestock mortality (above 51 percent).

[14] Variables on district-level infrastructure refer to 2012, when the first survey interview was conducted. We thus have to assume that conditions were constant between 2009 and 2012.

4.6 Empirical Results

4.6.1 Testing for Balance in Covariates

First of all, we explore the balance in covariates across treated households (who purchased IBLI in 2009) and non-treated households (who did not purchase IBLI in 2009). Table 4.4 displays mean values and tests on differences in means. Most importantly, treated and control sample households do not differ significantly in their pre-shock herd size. The average number of livestock in 2009 was 349 animals among treated herders and 309 animals among control herders. During the 2009/10 dzud, treated herders lost 130 animals, compared to 143 animals lost among control herders. Again, this difference in losses caused by the shock is not statistically significant for treated and control herders. Thus, it is reassuring that treated and control households exhibit similar characteristics in the key covariates in our analysis. This also holds for characteristics of the local ecological zone, which correlate strongly with the long-term dzud risk. Similarly, treated and control households share very similar levels of education and age of the head of household, as well as having a comparable share of female breeding stock among their animals in 2012. Treated and control households live in districts that share similar characteristics in transport infrastructure.

However, several other covariates exhibit statistically significant differences across treated and control households. For instance, treated households have a significantly higher likelihood to live in rural areas, they judge their relative subjective wellbeing in 2009 to be higher, and they are more risk averse than non-treated herders. Treated households tend to live in areas that exhibit, on average, a significantly lower livestock mortality in 2010. Furthermore, treated households are significantly less likely to know the governor in their sub-district very well and they have significantly more economically active household members. To conclude, assignment to treatment is not ignorable and without controlling for selection into purchasing IBLI in 2009, estimated effects of receiving IBLI indemnity payments would be biased.

4.6.2 The Effect of IBLI Payments on Recovery

Next, we employ the bias-corrected matching estimator to assess the impact of IBLI indemnity payments in 2010 on post-disaster livestock recovery. Table 5 shows average treatment effects for eight different outcome variables. Panel A, B and C display results when using the minimal, core and maximal set of covariates, respectively.

Results show an overall positive effect of IBLI indemnity payments on post-disaster livestock recovery after controlling for selection based on observables. Herders who purchased IBLI in 2009 and received indemnity payments in autumn 2010 have a larger herd size in 2011, 2012, and 2013 compared to herders who did not purchase IBLI (table 4.5, columns 1-3). The treatment effect is strongest for livestock holdings in 2012, with all three sets of covariates yielding statistically significant results. For livestock holdings in 2013, the treatment effect is only marginally statistically significant at the 13 percent level for the core set of covariates (panel B), while the treatment effect is no longer significant for the minimal set of covariates (panel A) and the maximal set of covariates (panel C). For livestock holdings in 2014 (column 4), the treatment effect is still positive, but no longer statistically significant for any of the three sets of covariates.

The magnitude of the treatment effect is relatively large: In 2011, treated households own on average 14 to 15 percent more livestock than control households; in 2012, they own between 20 percent and 24 percent more livestock; and in 2013, they own about 16 percent more livestock.[15] This corresponds to a difference in herd size between treated and control herders of about 20, 32, and 27 animals in 2011, 2012, and 2013; this is equals to an additional livestock wealth of the treated households between 1,150,000 and 1,173,000 MNT (between 780 and 1,173 US$). The positive effect of indemnity payments appears to attenuate three years after the shock.

These results are confirmed when focusing on the cumulative annual livestock growth rates as outcome variables (table 4.5, columns 5-8). Again, receiving IBLI indemnity payments in 2010 has a positive and significant effect on post-disaster livestock dynamics through 2012 and 2013. This finding is strongest for the core set of covariates (panel B), while the minimal and maximal set of covariates (panels A and C) only yield significant results for the period 2009-2012. On average, treated herders have an annual livestock growth rate that is about 4 to 5 percentage points higher compared to control households when considering the 2009-2012 period; and for the entire 2009-2013 period about 3 percentage points higher.

[15] Since the outcome is log transformed, the magnitude of the treatment effect is calculated as the exponentiated value of the estimated coefficient. We only consider coefficients that are at least marginally significant.

As a refinement, we investigate whether the positive effect is homogenous across different levels of indemnity payments received. We exploit information on the amount of indemnity payments received in 2010 as self-reported by respondents. Treated households in our sample reported receiving indemnity payments between 38,000 MNT (28 US$) and 1,300,000 MNT (974 US$), with the average being 268,000 MNT (201 US$). This is considerably lower than figures from official PIU records, which report that households in Uvs province received on average 416,000 MNT (312 US$) as indemnity payments in 2010 (PIU 2012). Thus, underreporting appears likely among the survey households. Therefore, we opted for distinguishing between three doses of treatment: receiving no indemnity payments (non-treated households), receiving indemnity payments below the 25th percentile (up to 120,000 MNT (90 US$)), and receiving indemnity payments above the 25th percentile. We focus on the recovery period until 2012, which provided the most robust results in the baseline estimation. Results in table 4.6 show that the estimated coefficients of both high and low doses of treatment are at least marginally significant. This holds when both using livestock holdings in 2012 and cumulative livestock growth rates in the period 2009-2012 as outcome variables. In both estimates, the estimated coefficients of high doses of treatment are larger in magnitude (although the difference is not statistically significant), suggesting that the positive effect of treatment is more pronounced for treated households that received high indemnity payments. This could indicate that there might be a minimum insurance coverage needed for positive effects to unfold.

4.6.3 Robustness Tests

We conduct various tests on the sensitivity of results to model assumptions. First, the estimated treatment effect derived from the matching estimator is sensitive to the number of matches chosen for each unit (Abadie et al. 2004). The bias due to inexact matches is likely to decrease with a larger set of matches (Abadie and Imbens 2006, p. 240). While we use four matches in our baseline estimations, table 4.7 shows results when using two and six observations for each match. Results are not sensitive to the number of matches. The magnitude of the estimated treatment effect slightly increases when a larger number of matches is used.

Second, we employ two alternative propensity score estimators that correct for self-selection into treatment based on observables (see table 4.8): the inverse probability-weighted regression adjustment (IPWRA) (panel A) and the augmented inverse-probability weighting (AIPW) (panel B). Both estimators perform two separate regressions on the treatment status and the outcome variable. The estimators are double-robust, as the estimated coefficients are robust to misspecifications in one of the two regressions (Wooldridge 2007; 2010). Results confirm the positive and significant effect of IBLI indemnity payments on post-disaster recovery until 2012. Overall, the magnitude of the treatment effect is slightly smaller compared to results from the matching estimator. The average treatment effect on livestock holdings in 2012 is 0.19 (IPWRA) and 0.21 (AIPW), compared to an estimate of 0.24 obtained with the bias-corrected matching estimator.

Third, we account for possible changes in herd composition. All results presented so far rely on total herd size, treating animals of different species as equal. Yet, the recovery of livestock losses after the dzud is also influenced by natural reproduction rates, which vary across species and are highest for sheep and goats. There is no evidence that treated herders had a higher share of small animals in their herd, which would have explained their faster recovery: the share of sheep and goats in the herd in 2009, 2011, 2012, and 2014 is not statistically different between treated and control households. In 2013, treated herders even had a slightly lower percentage of sheep and goats. Table 4.9 displays estimates of the main specification using the bias-corrected matching estimator in which now all livestock holdings are transformed into horse units (called bod units in Mongolian), the conversion rate commonly used in Mongolia. Again, results support our main findings: even when accounting for pre-shock and post-shock herd composition, receiving IBLI indemnity payments helped households recover from dzud losses.

Fourth, we explore the robustness of the retrospective livestock data. Recall that all results presented so far are based on information on pre-shock livestock holdings and livestock losses caused by the dzud recorded in the first panel wave. Table 4.10, panel A shows results when we instead employ retrospective livestock data recorded in the third panel wave. Panel B shows estimates obtained when using retrospective information only if reported by the head of household (in either the first or third panel wave). All main results hold when using these alternative retrospectively recorded measures. When using the retrospective livestock holdings as reported by the head of household, we even obtain a marginally significant and sizeable treatment effect for livestock holdings in 2014. This test illustrates the reliability of the retrospectively recorded information on past livestock holdings.

Lastly, we test whether treatment influenced households' livestock holdings immediately after the dzud, calculated as the difference between livestock holdings in 2009 and reported livestock losses due to dzud. If the estimated effects of treatment had a significant effect on this outcome, this would point toward unobservable factors at play that influence households' shock exposure and that we do not capture well with observed covariates. In line with expectations, results displayed in table 4.11 from the bias-corrected matching estimator, IPWRA, and AIPW show that IBLI indemnity payments do not have a significant effect on livestock holdings immediately after the dzud.

4.6.4 Unravelling the Channels

Lastly, we explore through what channels IBLI indemnity payments may have helped treated households recover faster. Recall that sample households that had purchased IBLI in 2009 received on average indemnity payments of 268,000 MNT (201 US$) in 2010. Although underreporting seems very likely, this amount is not impressively large. For instance, this amount would have bought nine female goats, seven female sheep, one female horse, one cow, or one female camel at prices prevailing in Uvs province in 2010 (NSO 2011). It represents merely 7 percent of the yearly cash income from herding that households earned in 2012, when climatic conditions were particularly favorable for livestock activities. However, when asked to assess how helpful the indemnity payments were, most treated households in our sample were very satisfied. Of the 59 sample households that had purchased IBLI in 2009, 44 households (75 percent) indicated that they found the indemnity payments helpful to manage the consequences of the dzud, compared to 14 households (24 percent) indicating that they found the indemnity payments either too small or coming too late. In fact, the majority of households that received IBLI indemnity payments in 2010 continued purchasing IBLI in the post-dzud period: Of the 59 treated households, 37, 37, and 14 households purchased IBLI again in 2012, 2013, and 2014.

Descriptive statistics indicate that treated sample households used IBLI indemnity payments mostly to cover household expenses. A large share of treated sample households – 43 households (73 percent) – used the indemnity payments received in 2010 to buy food and other household necessities. Eight households (14 percent) used the indemnity payments to cover education and health expenses. Only 13 households (22 percent) reported using the indemnity payments for investments in livestock activities, such as buying livestock fodder and improving shelters. Surprisingly, none of the treated households reported using the indemnity payments to restock the herd. Interestingly, nine households (15 percent) used the indemnity payments to pay back a loan.

The faster asset recovery of insured households also appears to be a result of different shock coping strategies applied. Table 4.12 presents results from the bias-corrected matching estimator on the effect of treatment on the usage of five different shock coping strategies. Treated households were significantly less likely forced to sell livestock during or after dzud (when prices were low) compared to non-treated households. Treated households were also less likely to move animals in the midst of the dzud winter, a common informal coping strategy. Furthermore, treated households were significantly more likely to borrow money during the dzud. The magnitude of this effect is very large. Through purchasing IBLI, insured herders became customers of local commercial banks. These banks offered credits at discounted interest rates to IBLI customers with the value of a household's livestock holdings as documented on the IBLI policy serving as collateral. Thus, this evidence suggests that IBLI indemnity payments helped treated households to relieve credit constraints during dzud and to smooth their productive assets.

Complementary qualitative in-depth interviews with a small sub-sample of herding households, local governors and insurance agents confirm this interpretation.[16] Several herders pointed out that it was not just livestock losses during the dzud months, but also the slaughtering of animals for meat consumption that made a difference for herd growth after the dzud:

> *Authors: What are the factors that helped you recover from the dzud so fast?*
>
> *Herder: I think our hard work accounts for the most part. By preparing winter fodder very well, we made sure that as many livestock as possible survived the winter.*
>
> *Authors: So you did not purchase any livestock? Received no aid?*
>
> *Herder: No, nothing at all. Having only one child, we consumed less than other households, which mainly accounts for our fast recovery.*

Another herder stressed how cash inflows can help herders avoid selling animals:

> *Herder: Money is the most helpful thing after dzud. This money can help us to buy primary consumption goods.*

Other herders would rather use cash inflows to restock the herd with new animals:

[16] We interviewed ten herders, five sub-district governors, one agricultural officer, three IBLI project officers, the manager of a private insurance company selling IBLI, and three insurance agents in western Mongolia in June 2014. The aim of the qualitative interviews was to understand better the role IBLI played in herders' post-shock recovery. All interviews were recorded, transcribed and then translated into English.

Authors: If you had gotten some cash after the dzud, would it have helped you to recover?

Herder: Yes. To a certain extent, we could have invested in purchasing female livestock from within the region to nurture faster growth.

It appears that herders have a preference to purchase new livestock in spring when prices are low:

Authors: When is it most ideal to buy livestock after dzud?

Herder: Generally, in spring, in May.

Authors: So shortly after dzud. Aren't the animals then very weak?

Herder: Livestock are cheaper when they are thin.

Authors: Do you buy the thinner ones to strengthen them?

Herder: Exactly. We buy livestock thin and strengthen them until autumn.

Other herders stated they prefer purchasing livestock in spring as this would allow the new livestock to adapt to new pastures before the winter starts. Thus, it appears as if the timing of IBLI indemnity payments (starting in August) was less suited to restock the herd in 2010. Rather, treated households may have used a credit to purchase livestock during or shortly after the dzud and repaid the loan with the IBLI indemnity payments.

To conclude, treated households appear to have benefitted from liquidity obtained from both IBLI indemnity payments and from relieved credit constraints. It seems that IBLI helped treated households recover faster by allowing them to avoid slaughtering and selling animals to pay for household expenses as well as by purchasing livestock to restock their herds.

4.7 Conclusion

Index insurance is praised as a powerful tool that supports smallholder farmers and herders in developing countries in managing weather risk. Yet, there is scant empirical evidence to date on the actual benefits of index insurance for agricultural households. Our study is among the first that empirically investigates the *ex-post* impacts of indemnity payments from index insurance after a shock occurred. Our focus is on IBLI Mongolia, which is a fully commercial product available at the national level since 2012. We analyze the effect of IBLI indemnity payments after a once-in-50-year winter disaster struck Mongolia in 2009/10. This event caused the worst livestock losses ever recorded in a single winter. Our analysis tests if IBLI indemnity payments helped insured herders to recover their herd size faster than non-insured herders. The database for our analysis is three waves of a household panel survey implemented in western Mongolia. One particular feature of the survey is that it asks households retrospectively about the purchase of IBLI before the winter disaster, their shock exposure, coping strategies applied as well as livestock holdings at different points in time. We employ the bias-corrected matching estimator to account for selection into purchasing IBLI in 2009, just before the disaster occurred. Our empirical strategy exploits further the phasing-in of the IBLI scheme. This helps us to exclude the possibility of spillover effects and minimizing the potential bias stemming from unobserved characteristics influencing both the selection into treatment and outcome variables.

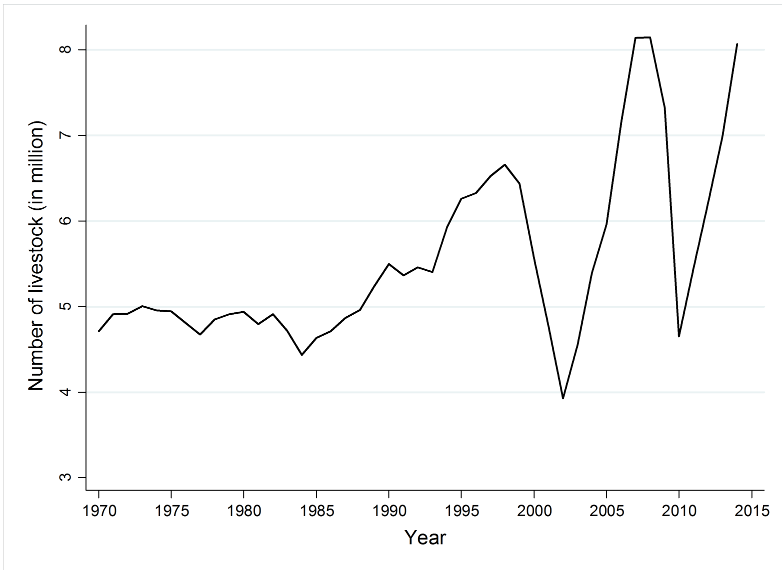
Pastoralist households purchasing IBLI before the shock recovered faster from shock-induced asset losses than comparable non-insured households. We find a significant, positive and economically large effect of IBLI indemnity payments on herd size both one and two years after the shock. In the medium term – three and four years after the shock – the effect is still visible but narrowing. These findings, obtained with the bias-corrected matching estimator, hold both when using livestock holdings in the post-shock period and cumulative growth rates in livestock. Results are also robust to varying the number of matches per observation, the choice of covariates, and to the usage of alternative double robust estimators. Also, we can exclude the possibility that the effect is driven by a change in herd composition toward smaller animals with higher reproduction rates among treated households.

An analysis of shock coping strategies as well as complementary qualitative interviews conducted in the field suggest that herders benefit from IBLI indemnity payments through two channels: On the one hand, indemnity payments are used to cover expenses for food, education, and health. Herders can thus avoid selling and slaughtering animals and smooth their productive asset base. On the other hand, IBLI appears to have relieved households from credit constraints, which may have been used to purchase new livestock. Access to credit appears to be a positive side-effect of IBLI on rural financial markets.

Our analysis is restricted to some limitations and shortcomings. First, our analysis rests on the – untestable – assumption that we capture the voluntary purchase decision of IBLI with observed covariates, with no unobservable factors remaining. Second, all treated households in our sample stayed in the herding economy in the aftermath of the shock. Hence, it is not possible to draw conclusions on whether IBLI helped households to avoid dropping out of the herding economy. Finally, the small number of treated sample households does not allow for a more detailed analysis of heterogeneous treatment effects.

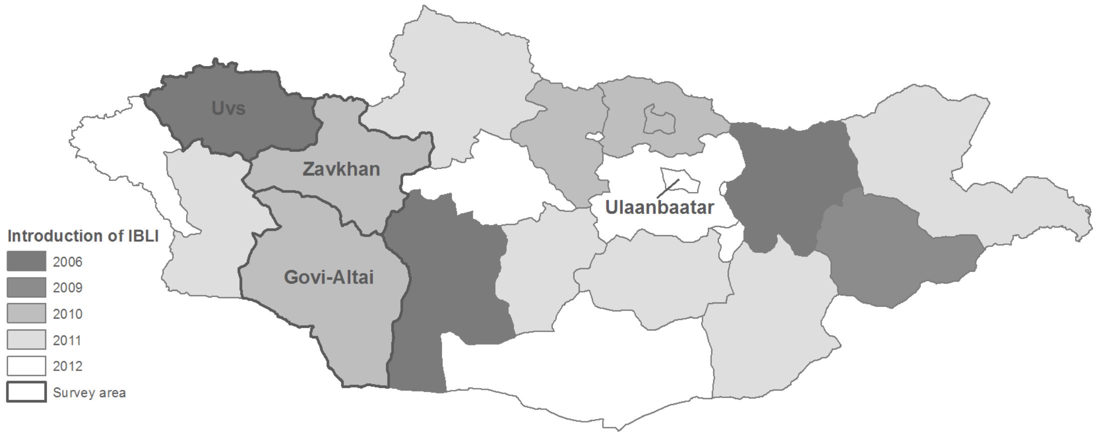
Tables and Figures

Figure 4.1: Livestock development in western Mongolia, 1970-2014



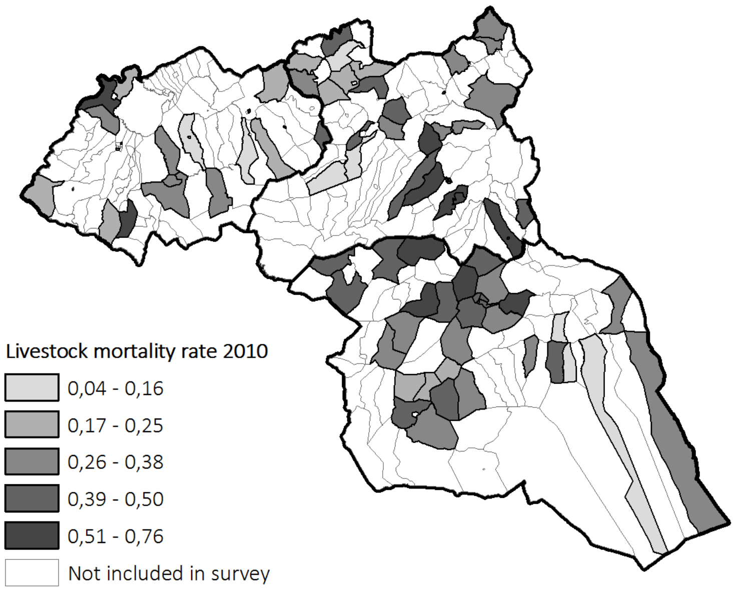
Notes: Livestock include camel, cattle, horse, sheep, and goat. Data shown for the provinces Uvs, Zavkhan, and Govi-Altai. Source: Mongolia Livestock Census.

Figure 4.2: Map of Mongolia, showing the year in which IBLI was introduced



Notes: The three provinces where the household survey was implemented are bold-rimmed. Adapted from PIU (2012).

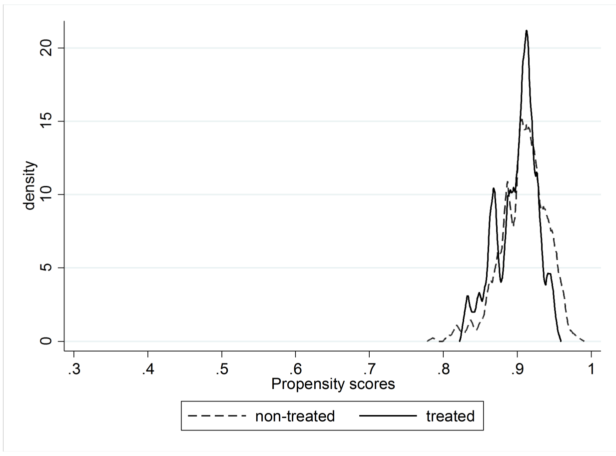
Figure 4.3: Map of the survey area, showing livestock mortality in 2010 per sub-district



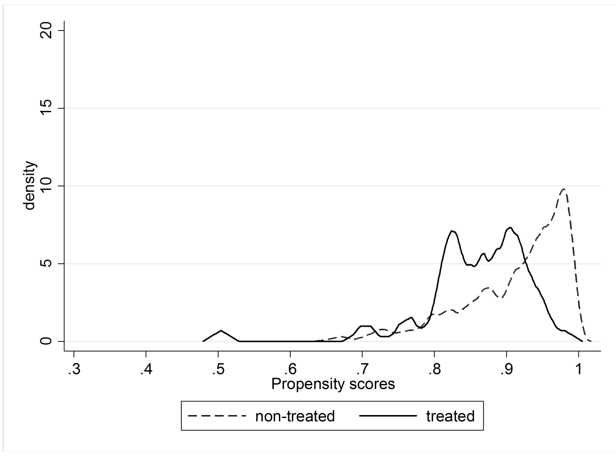
Notes: The map shows the three provinces of Uvs, Zavkhan, and Govi-Altai where the household survey was implemented. Source: Mongolia Livestock Census.

Figure 4.4: Overlap in covariates across treated and control households

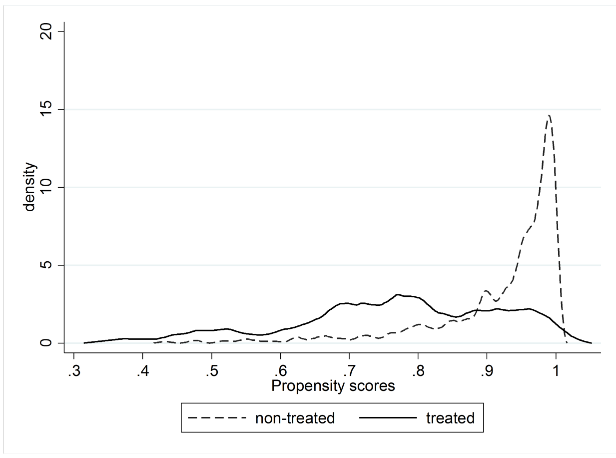
(a) Minimal set of covariates



(b) Core set of covariates



(c) Maximal set of covariates



Notes: The figures show the overlap in the propensity scores of covariates of treated and non-treated households. See table 4.5 for details on the definition of the set of covariates. The propensity scores are estimated using the augmented-inverse probability weighting estimation with number of livestock (log) in 2012 as outcome variable. Source: Coping with Shocks in Mongolia Household Panel Survey and Mongolia Livestock Census.

Table 4.1: Summary statistics

	mean	sd	min	max	observations
outcome variables					
number of livestock 2011	194.48	176.51	2	1,563	642
number of livestock 2012	199.80	183.23	3	1,613	642
number of livestock 2013	242.59	220.62	5	1,609	608
number of livestock 2014	277.85	255.06	4	1,867	581
cum. livestock growth rate, 2009-2011	-0.21	0.28	-0.88	0.75	642
cum. livestock growth rate, 2009-2012	-0.15	0.20	-0.70	0.58	642
cum. livestock growth rate, 2009-2013	-0.09	0.17	-0.69	0.48	608
cum. livestock growth rate, 2009-2014	-0.05	0.15	-0.47	0.37	581
coping strategy: sold livestock	0.17	0.37	0	1	642
coping strategy: moved livestock during dzud	0.36	0.48	0	1	642
coping strategy: borrowed money	0.32	0.47	0	1	642
coping strategy: organized add. labor for herding	0.21	0.41	0	1	642
coping strategy: built shelter or fences for livestock	0.09	0.29	0	1	642
treatment variables					
purchased IBLI in 2009	0.09	0.29	0	1	642
value of IBLI payouts household received in 2010 (in thousand MNT)	268.29	245.92	38	1,300	58
household head controls					
no education	0.13	0.34	0	1	642
primary education	0.58	0.49	0	1	642
secondary education	0.29	0.45	0	1	642
age	44.66	12.39	19	87	642
risk preference (0=risk averse, 10=risk loving)	4.24	3.39	0	10	642
knows the sub-district governor very well	0.50	0.50	0	1	642
household controls					
number of livestock in 2009	312.84	236.09	10	1,800	642
number of livestock lost due to 2009/2010 dzud	142.15	135.95	1	950	642
percent of breeding stock	0.37	0.09	0	0.90	642
relative subjective economic wellbeing in 2009 (0=among the poorest, 10=among the richest)	5.77	1.50	1	10	642
number of economically active members	2.08	1.05	0	7	642
location is rural	0.62	0.48	0	1	642
sub-district controls					
livestock mortality in 2010	0.37	0.14	0.05	0.76	642
district controls					
ecological zone is mountain steppe	0.26	0.44	0	1	642
ecological zone is forest steppe	0.13	0.34	0	1	642
ecological zone is grass steppe	0.29	0.45	0	1	642
ecological zone is desert steppe/desert	0.33	0.47	0	1	642
cellphone coverage (1=in few areas; 4=in all areas)	2.74	0.86	1	4	642
number of transport options to provincial center	1.51	0.86	0	3	642

Source: Coping with Shocks in Mongolia Household Panel Survey and Mongolia Livestock Census.

Table 4.2: Determinants of purchasing IBLI in 2009 (probit)

	(1)	(2)
outcome variable	purchased IBLI	purchased IBLI
household head controls		
primary education	-0.029	0.101
	(0.93)	(0.79)
secondary education	0.203	0.479
	(0.55)	(0.26)
age	0.006	0.009
	(0.43)	(0.36)
risk preference	0.064	0.064
	(0.12)	(0.20)
household controls		
number of livestock in 2009 (logs)	0.330**	0.608***
	(0.03)	(0.00)
relative subjective economic wellbeing in 2009	0.039	0.074
	(0.55)	(0.34)
location is rural	-0.133	0.016
	(0.61)	(0.95)
district controls		
ecological zone is mountain steppe	-0.165	
	(0.51)	
ecological zone is forest steppe	0.737**	
	(0.04)	
ecological zone is grass steppe	0.323	
	(0.22)	
cellphone coverage	-0.162	
	(0.11)	
number of transportation options to provincial center	-0.043	
	(0.79)	
constant	-2.734***	-3.597***
	(0.00)	(0.00)
district fixed effects	NO	YES
observations	263	188

Notes: P-values are reported in brackets with * significant at 10%; ** significant at 5%; *** significant at 1%. The sample comprises all herders in the province of Uvs where IBLI was available in 2009. Source: Coping with Shocks in Mongolia Household Panel Survey and Mongolia Livestock Census.

Table 4.3: Determinants of household livestock losses due to the 2009/10 dzud (OLS)

outcome variable	number of livestock lost due to 2009/10 dzud (in logs)
purchased IBLI in 2009	-0.082
	(0.91)
household head controls	
primary education	-0.032
	(0.69)
secondary education	-0.096
	(0.30)
age	0.002
	(0.36)
risk preference	-0.021***
	(0.00)
household controls	
number of livestock in 2009 (in logs)	0.902***
	(0.00)
relative subjective economic wellbeing in 2009	0.007
	(0.73)
location is rural	-0.171
	(0.10)
sub-district controls	
livestock mortality in 2010	0.458
	(0.21)
constant	-0.345
	(0.45)
district fixed effects	YES
R2	0.587
observations	642

Notes: P-values are reported in brackets with * significant at 10%; ** significant at 5%; *** significant at 1%.
Source: Coping with Shocks in Mongolia Household Panel Survey and Mongolia Livestock Census.

Table 4.4: Comparison of treated and non-treated households

	mean		p-values
	treated households N=59	non-treated households N=583	
	(1)	(2)	(3)
household head controls			
no education	0.10	0.13	0.51
primary education	0.56	0.58	0.72
secondary education	0.34	0.28	0.38
age	45.08	44.61	0.78
risk preference	2.59	4.41	0.00***
knows the sub-district governor very well	0.20	0.53	0.00***
household controls			
number of livestock in 2009	349.15	309.17	0.22
number of livestock lost due to 2009/2010 dzud	130.07	143.38	0.47
percent of breeding stock	0.38	0.37	0.49
relative subjective economic wellbeing in 2009	6.10	5.74	0.08*
number of economically active members	2.56	2.03	0.00***
location is rural	0.76	0.61	0.02**
sub-district controls			
livestock mortality in 2010	0.31	0.37	0.00***
district controls			
ecological zone is mountain steppe	0.20	0.26	0.32
ecological zone is forest steppe	0.15	0.13	0.58
ecological zone is grass steppe	0.27	0.29	0.78
ecological zone is desert steppe/desert	0.37	0.32	0.43
cellphone coverage	2.58	2.75	0.13
number of transport options to provincial center	1.41	1.52	0.34

Notes: Colum 3 shows p-values on tests on differences in means between treated and non-treated households. T-tests are used for continuous variables, chi-square tests for non-continuous variables with * significant at 10%; ** significant at 5%; *** significant at 1%. Source: Coping with Shocks in Mongolia Household Panel Survey and Mongolia Livestock Census.

Table 4.5: Impact of IBLI indemnity payments on recovery from the 2009/10 dzud (bias-corrected matching estimator, baseline results)

outcome variables	number of livestock (in logs)				cum. livestock growth rate			
	2011	2012	2013	2014	2009-2011	2009-2012	2009-2013	2009-2014
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treatment variable								
panel A: minimal set of covariates								
purchased IBLI in 2009	0.143**	0.197**	0.121	0.093	0.022	0.040**	0.021	0.012
	(0.04)	(0.01)	(0.20)	(0.33)	(0.36)	(0.05)	(0.30)	(0.48)
panel B: core set of covariates								
purchased IBLI in 2009	0.151	0.241**	0.156	0.108	0.026	0.052***	0.029*	0.015
	(0.10)	(0.02)	(0.13)	(0.33)	(0.33)	(0.00)	(0.08)	(0.33)
panel C: maximal set of covariates								
purchased IBLI in 2009	0.121	0.223**	0.096	0.072	0.016	0.048**	0.017	0.008
	(0.23)	(0.04)	(0.39)	(0.53)	(0.62)	(0.04)	(0.38)	(0.62)
mean outcome of control households	4.747	4.741	4.940	5.085	-0.211	-0.157	-0.087	-0.042
observations	642	642	608	581	642	642	608	581

Notes: Displayed is the estimated coefficient of treatment (ATE) obtained from 24 separate estimations. In Panel A, the minimal set of covariates includes number of livestock in 2009, number of livestock lost during dzud, and ecological zone. In Panel B, the core set of covariates additionally includes head of household controls (age, education, risk preference), household controls (relative subjective economic wellbeing in 2009, location of residence), sub-district controls (livestock mortality in 2010), and district controls (cellphone coverage, transportation). In Panel C, the maximal set of covariates additionally includes the relationship to the local governor, percentage of breeding stock, and number of economically active household members. The mean outcome of control households reported is the geometric mean for outcomes expressed in logs and the arithmetic mean otherwise. Four matches are used for every observation. P-values are reported in brackets with * significant at 10%; ** significant at 5%; *** significant at 1%. Source: Coping with Shocks in Mongolia Household Panel Survey and Mongolia Livestock Census.

Table 4.6: Impact of doses of treatment on recovery from the 2009/10 dzud (OLS with propensity score weighting)

outcome variables	livestock in 2012 (in logs)	cum. livestock growth rate 2009-2012
	(1)	(2)
treatment variables		
received low indemnity payments in 2010 (<90US$)	0.139**	0.026*
	(0.01)	(0.08)
received high indemnity payments in 2010 (>90US$)	0.167**	0.030*
	(0.01)	(0.10)
core set of covariates	YES	YES
mean outcome of control households	4.743	-0.157
observations	642	642

Notes: Displayed is the estimated coefficient of treatment (ATE) obtained from 2 separate weighted estimations with the core set of covariates included (see table 4.5 for details). The baseline category is not receiving indemnity payments. P-values are reported in brackets with * significant at 10%; ** significant at 5%; *** significant at 1%. Source: Coping with Shocks in Mongolia Household Panel Survey and Mongolia Livestock Census.

Table 4.7: Robustness test: Results from bias-corrected matching estimator with varying number of matches per observation)

outcome variables	number of livestock (in logs)				cum. livestock growth rate			
	2011	2012	2013	2014	2009-2011	2009-2012	2009-2013	2009-2014
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treatment variable								
panel A: two matches per observation								
purchased IBLI in 2009	0.145	0.230**	0.161	0.078	0.024	0.049**	0.030*	0.010
	(0.12)	(0.03)	(0.12)	(0.49)	(0.38)	(0.02)	(0.08)	(0.54)
panel B: six matches per observation								
purchased IBLI in 2009	0.162*	0.253**	0.166	0.125	0.030	0.056***	0.031*	0.018
	(0.08)	(0.01)	(0.11)	(0.27)	(0.27)	(0.00)	(0.06)	(0.25)
core set of covariates	YES	YES	YES	YES	YES	YES	YES	YES
mean outcome of control households	4.747	4.741	4.940	5.085	-0.211	-0.157	-0.087	-0.042
observations	642	642	608	581	642	642	608	581

Notes: Displayed is the estimated coefficient of treatment (ATE) obtained from 16 separate estimations with the core set of covariates included (see table 4.5 for details). P-values are reported in brackets with * significant at 10%; ** significant at 5%;*** significant at 1%.
Source: Coping with Shocks in Mongolia Household Panel Survey and Mongolia Livestock Census.

Table 4.8: Robustness test: Results from double robust estimators

outcome variables	number of livestock (in logs)				cum. livestock growth rate			
	2011	2012	2013	2014	2009-2011	2009-2012	2009-2013	2009-2014
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treatment variable								
panel A: inverse probability-weighted regression adjustment (IPWRA)								
purchased IBLI in 2009	0.059	0.192***	0.003	0.078	-0.009	0.039**	-0.005	0.009
	(0.25)	(0.00)	(0.97)	(0.40)	(0.63)	(0.01)	(0.81)	(0.60)
panel B: augmented inverse-probability weighting (AIPW)								
purchased IBLI in 2009	0.152	0.209**	-0.040	0.070	-0.012	0.044***	-0.015	0.008
	(0.26)	(0.00)	(0.63)	(0.43)	(0.46)	(0.00)	(0.44)	(0.65)
core set of covariates	YES	YES	YES	YES	YES	YES	YES	YES
mean outcome of control households	4.747	4.741	4.940	5.085	-0.211	-0.157	-0.087	-0.042
observations	642	642	608	581	642	642	608	581

Notes: Displayed is the estimated coefficient of treatment (ATE) obtained from 16 separate estimations with the core set of covariates included (see table 4.5 for details). P-values are reported in brackets with * significant at 10%; ** significant at 5%; *** significant at 1%.
Source: Coping with Shocks in Mongolia Household Panel Survey and Mongolia Livestock Census.

Table 4.9: Impact of IBLI indemnity payments on recovery from the 2009/10 dzud based on livestock in bod units (bias-corrected matching estimator)

outcome variables	number of livestock in bod units (in log)			
	2011	2012	2013	2014
	(1)	(2)	(3)	(4)
treatment variable				
purchased IBLI in 2009	0.237**	0.385***	0.263**	0.251*
	(0.02)	(0.00)	(0.03)	(0.06)
core set of covariates	YES	YES	YES	YES
mean outcome of control households	3.041	3.033	3.263	3.449
observations	430	430	420	416

Notes: Displayed is the estimated coefficient of treatment (ATE) obtained from 4 separate weighted estimations with the core set of covariates included (see table 4.5 for details) except that livestock holdings in 2009 and livestock losses due to dzud are now expressed in bod units. P-values are reported in brackets with * significant at 10%; ** significant at 5%; *** significant at 1%. Source: Coping with Shocks in Mongolia Household Panel Survey and Mongolia Livestock Census.

Table 4.10: Robustness test: Using retrospective information on pre-shock livestock holdings and livestock losses from the third panel wave (bias-corrected matching estimator)

outcome variables	number of livestock (in logs)				cum. livestock growth rate			
	2011	2012	2013	2014	2009-2011	2009-2012	2009-2013	2009-2014
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treatment variable								
panel A: retrospective information on livestock recorded in third panel wave								
purchased IBLI in 2009	0.195	0.290**	0.130	0.057	0.053	0.071**	0.026	0.006
	(0.13)	(0.03)	(0.39)	(0.70)	(0.18)	(0.02)	(0.34)	(0.77)
core set of covariates	YES	YES	YES	YES	YES	YES	YES	YES
mean outcome of control households	4.769	4.758	4.938	5.114	-0.211	-0.161	-0.092	-0.048
observations	528	528	503	494	528	528	503	494
panel B: retrospective information on livestock reported by head of household (first or third panel wave)								
purchased IBLI in 2009	0.137	0.254***	0.176**	0.144	0.024	0.058***	0.035**	0.022
	(0.10)	(0.00)	(0.04)	(0.15)	(0.40)	(0.00)	(0.04)	(0.19)
core set of control households	YES	YES	YES	YES	YES	YES	YES	YES
mean outcome of control households	4.777	4.761	4.937	5.071	-0.198	-0.150	-0.083	-0.047
observations	566	566	538	522	566	566	538	522

Notes: Displayed is the estimated coefficient of treatment (ATE) obtained from 16 separate estimations with the core set of covariates included (see table 4.5 for details). P-values are reported in brackets with * significant at 10%; ** significant at 5%; *** significant at 1%. Source: Coping with Shocks in Mongolia Household Panel Survey and Mongolia Livestock Census.

Table 4.11: Robustness test: Impact of IBLI on livestock holdings in 2010

outcome variable	number of livestock in 2010 (logs)		
	bias-corrected matching estimator	inverse probability-weighted regression (IPWRA)	augmented inverse-probability weighting (AIPW)
	(1)	(2)	(3)
treatment variables			
purchased IBLI in 2009	0.082	0.034	0.028
	(0.49)	(0.46)	(0.49)
core set of covariates	YES	YES	YES
mean outcome of control households	4.759	4.759	4.759
observations	639	639	639

Notes: Displayed is the estimated coefficient of treatment (ATE) obtained from 3 separate estimations with the core set of covariates included (see table 4.5 for details). P-values are reported in brackets with * significant at 10%; ** significant at 5%; *** significant at 1%. Source: Coping with Shocks in Mongolia Household Panel Survey and Mongolia Livestock Census.

Table 4.12: Impact of IBLI indemnity payments on household shock coping strategies (bias-corrected matching estimator, baseline results)

	coping strategies				
outcome variables	sold livestock	moved livestock during dzud	borrowed money	organized additional labor for herding	built shelter or fences for livestock
	(1)	(2)	(3)	(4)	(5)
treatment variable					
panel A: minimal set of covariates					
purchased IBLI in 2009	-0.104**	-0.216***	0.333***	0.008	-0.006
	(0.03)	(0.00)	(0.00)	(0.90)	(0.89)
panel B: core set of covariates					
purchased IBLI in 2009	-0.154***	-0.094*	0.346***	0.103	0.016
	(0.00)	(0.09)	(0.00)	(0.17)	(0.70)
panel C: maximal set of covariates					
purchased IBLI in 2009	-0.149***	-0.062	0.249***	0.156*	0.030
	(0.00)	(0.23)	(0.00)	(0.09)	(0.58)
mean outcome of control households	0.18	0.38	0.289	0.207	0.094
observations	642	642	642	642	642

Notes: Displayed is the estimated coefficient of treatment (ATE) obtained from 15 separate estimations based on different sets of covariates (see table 4.5 for details). P-values are reported in brackets with * significant at 10%; ** significant at 5%; *** significant at 1%. Source: Coping with Shocks in Mongolia Household Panel Survey and Mongolia Livestock Census.

References

Abadie, A., D. Drukker, J. L. Herr, and G. W. Imbens (2004). Implementing matching estimators for average treatment effects in Stata. *Stata Journal 4*, 290–311.

Abadie, A. and G. Imbens (2002). Simple and bias-corrected matching estimators for average treatment effects. NBER Working Paper 283, National Bureau of Economic Research, Cambridge, MA.

Abadie, A. and G. W. Imbens (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica 74*(1), 235–267.

Abadie, A. and G. W. Imbens (2011). Bias-corrected matching estimators for average treatment effects. *Journal of Business & Economic Statistics 29*(1), 1–11.

Alpizar, F., F. Carlsson, and O. Johansson-Stenman (2005). How much do we care about absolute versus relative income and consumption? *Journal of Economic Behavior & Organization 56*(3), 405–421.

Andersen, S., J. Fountain, G. W. Harrison, and E. E. Rutström (2014). Estimating subjective probabilities. *Journal of Risk and Uncertainty 48*(3), 207–229.

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

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

Anderson, L. R. and J. M. Mellor (2009). Are risk preferences stable? Comparing an experimental measure with a validated survey-based measure. *Journal of Risk and Uncertainty 39*(2), 137–160.

Ariely, D. and S. Jones (2008). *Predictably irrational*. Harper Collins New York.

Attanasio, O., A. Barr, J. C. Cardenas, G. Genicot, and C. Meghir (2012). Risk pooling, risk preferences, and social networks. *American Economic Journal: Applied Economics 4*(2), 134–167.

Barnett, B. J., C. B. Barrett, and J. R. Skees (2008). Poverty traps and index-based risk transfer products. *World Development 36*(10), 1766–1785.

Barnett, B. J. and O. Mahul (2007). Weather index insurance for agriculture and rural areas in lower-income countries. *American Journal of Agricultural Economics 89*(5), 1241–1247.

Barr, A., M. Dekker, and M. Fafchamps (2012). Who shares risk with whom under different enforcement mechanisms? *Economic Development and Cultural Change 60*(4), 677–706.

Batima, P. (2006). Climate change vulnerability and adaptation in the livestock sector of Mongolia. Final Report Submitted to Assessments of Impacts and Adaptations to Climate Change (AIACC), International START Secretariat, Washington DC.

Bhandari, P. (2004). Relative deprivation and migration in an agricultural setting of Nepal. *Population and Environment 25*(5), 475–499.

Biemer, P., R. Groves, L. Lyberg, N. Mathiowetz, and S. Sudman (Eds.) (2011). *Measurement errors in surveys*, Volume 173. Hoboken, New Jersey, US: John Wiley & Sons.

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

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

Binswanger-Mkhize, H. P. (2012). Is there too much hype about index-based agricultural insurance? *The Journal of Development Studies 48*(2), 187–200.

Birdsall, N. and J. L. Londono (1997). Asset inequality matters: An assessment of the World Bank's approach to poverty reduction. *The American Economic Review 87*(2), 32–37.

Bossert, W., S. Chakravarty, and C. D'Ambrosio (2013). Multidimensional poverty and material deprivation with discrete data. *Review of Income and Wealth 59*(1), 29–43.

Boyce, C., G. Brown, and S. Moore (2010). Money and happiness rank of income, not income, affects life satisfaction. *Psychological Science 21*(4), 471–475.

Brown, G., J. Gardner, A. Oswald, and J. Qian (2008). Does wage rank affect employees' well-being? *Industrial Relations: A Journal of Economy and Society 47*(3), 355–389.

Brück, T., D. Esenaliev, A. Kröger, A. Kudebayeva, B. Mirkasimov, and S. Steiner (2014). Household survey data for research on well-being and behavior in Central Asia. *Journal of Comparative Economics 42*(3), 819 – 835.

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

Callen, M., M. Isaqzadeh, J. D. Long, and C. Sprenger (2014). Violence and risk preference: Experimental evidence from Afghanistan. *The American Economic Review 104*(1), 123–148.

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

Carter, M. (2009). Innovations in insuring the poor: Intelligent design of index insurance for smallholder farmers and pastoralists. 2020 Vision Focus Brief 17(6), IFPRI, Washington D.C.

Carter, M., A. de Janvry, E. Sadoulet, and A. Sarris (2014). Index-based weather insurance for developing countries: A review of evidence and a set of propositions for up-scaling. Development Policies Working Paper 111, FERDI, Clermont-Ferrand.

Carter, M., G. Elabed, and E. Serfilippi (2015). Behavioral economic insights on index insurance design. *Agricultural Finance Review 75*(1), 8–18.

Carter, P. M. R. and C. B. Barrett (2006). The economics of poverty traps and persistent poverty: An asset-based approach. *The Journal of Development Studies 42*(2), 178–199.

Charness, G., U. Gneezy, and A. Imas (2013). Experimental methods: Eliciting risk preferences. *Journal of Economic Behavior & Organization 87*, 43–51.

Charness, G. and A. Viceisza (2012). Comprehension and risk elicitation in the field: Evidence from rural Senegal. Departmental Working paper, Department of Economics, University of California at Santa Barbara.

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

Clark, A. and C. Senik (2010). Who compares to whom? The anatomy of income comparisons in Europe. *Economic Journal 120*(544), 573–594.

Clarke, D. J. (2011). A theory of rational demand for index insurance. Discussion Paper 572, Department of Economics, University of Oxford.

Cole, S., D. Stein, and J. Tobacman (2014). Dynamics of demand for index insurance: Evidence from a long-run field experiment. *The American Economic Review 104*(5), 284–290.

Cole, S. A., X. Giné, and J. I. Vickery (2013). How does risk management influence production decisions? Evidence from a field experiment. World Bank Policy Research Working Paper 6546, World Bank, Washington D.C.

Crosetto, P. and A. Filippin (2013). A theoretical and experimental appraisal of five risk elicitation methods. Jena Economic Research Papers 2013-009, Jena.

D'Ambrosio, C. and J. Frick (2007). Income satisfaction and relative deprivation: An empirical link. *Social Indicators Research 81*(3), 497–519.

D'Ambrosio, C. and J. Frick (2012). Individual wellbeing in a dynamic perspective. *Economica 79*(314), 284–302.

Dasgupta, P. (1997). Nutritional status, the capacity for work, and poverty traps. *Journal of Econometrics 77*(1), 5–37.

Dave, C., C. C. Eckel, C. A. Johnson, and C. Rojas (2010). Eliciting risk preferences: When is simple better? *Journal of Risk and Uncertainty 41*(3), 219–243.

De Angelis, K. (2013). Index-based livestock insurance: The case of Mongolia. Inside Stories on climate compatible development, CDKN, London.

De Brauw, A. and P. Eozenou (2014). Measuring risk attitudes among Mozambican farmers. *Journal of Development Economics 111*, 61–74.

Deaton, A. (2001). Relative deprivation, inequality, and mortality. NBER Working Paper w8099, National Bureau of Economic Research, Cambridge, MA.

Dercon, S. (2005). *Insurance against poverty*. Oxford University Press.

Dercon, S. and L. Christiaensen (2011). Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia. *Journal of Development Economics 96*(2), 159–173.

Diener, E. (2009). *The science of well-being: The collected works of Ed Diener*, Volume 37. Springer Science & Business Media.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association 9*(3), 522–550.

Easterlin, R. (1995). Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior & Organization 27*(1), 35–47.

Easterlin, R. (2001). Income and happiness: Towards a unified theory. *Economic Journal 111*(473), 465–484.

Eibner, C. and W. Evans (2005). Relative deprivation, poor health habits, and mortality. *Journal of Human Resources 40*(3), 591–620.

Elabed, G. and M. R. Carter (2014). Ex-ante impacts of agricultural insurance: Evidence from a field experiment in Mali. Unpublished.

Fafchamps, M. (2003). *Rural poverty, risk and development*. Edward Elgar Publishing.

Fafchamps, M. and F. Shilpi (2008). Subjective welfare, isolation, and relative consumption. *Journal of Development Economics 86*(1), 43–60.

FAO (1982). Livestock production in tropical Africa. `http://www.fao.org/wairdocs/ilri/x5443e/x5443e04.htm`. Accessed: 2015.09.07.

Feibel, B. J. (2002). Calculating investment returns. In F. J. Fabozzi (Ed.), *Handbook of Finance*, pp. 35–66. Wiley Online Library.

Ferguson, N. and M. Michaelsen (2013). The legacy of conflict: Regional deprivation and school performance in Northern Ireland. Ruhr Economic Paper 419, Essen.

Fernández-Giménez, M. E., B. Batkhishig, B. Batbuyan, and T. Ulambayar (2015). Lessons from the dzud: Community-based rangeland management increases the adaptive capacity of Mongolian herders to winter disasters. *World Development 68*, 48–65.

Ferrer-i-Carbonell, A. (2005). Income and well-being: An empirical analysis of the comparison income effect. *Journal of Public Economics 89*(5), 997–1019.

Gaurav, S., S. Cole, and J. Tobacman (2011). Marketing complex financial products in emerging markets: Evidence from rainfall insurance in India. *Journal of Marketing Research 48*, 150–162.

Giesbert, L., S. Steiner, and M. Bendig (2011). Participation in micro life insurance and the use of other financial services in Ghana. *Journal of Risk and Insurance 78*(1), 7–35.

Giné, X., R. Townsend, and J. Vickery (2007). Statistical analysis of rainfall insurance payouts in southern India. *American Journal of Agricultural Economics 89*(5), 1248–1254.

Giné, X., R. Townsend, and J. Vickery (2008). Patterns of rainfall insurance participation in rural India. *The World Bank Economic Review 22*(3), 539–566.

Giné, X. and D. Yang (2009). Insurance, credit, and technology adoption: Field experimental evidencefrom Malawi. *Journal of Development Economics 89*(1), 1–11.

Gneezy, U. and J. Potters (1997). An experiment on risk taking and evaluation periods. *The Quarterly Journal of Economics 112*(2), 631–645.

Goodland, A., D. Sheehy, and T. Shine (2009). Mongolia: Livestock sector study, volumne I - Synthesis report. World Bank, Washington D.C.

Greatrex, H., J. Hansen, S. Garvin, R. Diro, M. Le Guen, S. Blakeley, K. Rao, and D. Osgood (2015). Scaling up index insurance for smallholder farmers: Recent evidence and insights. CCAFS Report, CGIAR, Copenhagen.

Greene, W. (2010). *Econometric Analysis* (6th ed.). Englewood Cliffs, New Jersey, US: Prentice Hall.

Grimm, M., C. Guenard, and S. Mesple-Somps (2002). What has happened to the urban population in Côte d'Ivoire since the 1980s? An analysis of monetary poverty and deprivation over 15 years of household data. *World Development 30*(6), 1073–1095.

Groppo, V. and K. Schindler (2014). Extreme weather events and child height: Evidence from Mongolia. DIW Berlin Discussion Paper 1403, DIW Berlin.

Guo, S. and M. W. Fraser (2014). *Propensity score analysis: Statistical methods and applications*, Volume 11. Sage Publications.

Hanoch, Y., J. G. Johnson, and A. Wilke (2006). Domain specificity in experimental measures and participant recruitment an application to risk-taking behavior. *Psychological Science 17*(4), 300–304.

Harbaugh, W. T., K. Krause, and L. Vesterlund (2010). The fourfold pattern of risk attitudes in choice and pricing tasks. *The Economic Journal 120*(545), 595–611.

Hardeweg, B., L. Menkhoff, and H. Waibel (2013). Experimentally validated survey evidence on individual risk attitudes in rural Thailand. *Economic Development and Cultural Change 61*(4), 859–888.

Harless, D. W. and C. F. Camerer (1994). The predictive utility of generalized expected utility theories. *Econometrica 62*(6), 1251–1289.

Harrison, G. W. (2008). Maximum likelihood estimation of utility functions using Stata. Working Paper Number 06-12, Department of Economics, University of Central Florida.

Harrison, G. W., S. J. Humphrey, and A. Verschoor (2010). Choice under uncertainty: evidence from Ethiopia, India and Uganda. *The Economic Journal 120*(543), 80–104.

Harrison, G. W. and M. I. Lau (2014). Risk attitudes, sample selection and attrition in a longitudinal field experiment. CEAR Working Paper 2014-04, Robinson College of Business, Georgia State University.

Harrison, G. W., J. Martínez-Correa, and J. T. Swarthout (2012). Reduction of compound lotteries with objective probabilities: Theory and evidence. CEAR Working Paper 2012-05, Robinson College of Business, Georgia State University.

Harrison, G. W. and E. E. Rutström (2008). Risk aversion in the laboratory. In J. Cox and G. Harrison (Eds.), *Risk Aversion in Experiments*, Volume 12, Chapter 6, pp. 41–196. Emerald Group Publishing Limited.

Hartell, J. (2011). Developing an impact assessment strategy for the Mongolian Index-Based Livestock Insurance pilot. GlobalAgRisk, Lexington.

Henrich, J. and R. McElreath (2002). Are peasants risk-averse decision makers? *Current Anthropology 43*(1), 172–181.

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

Hey, J. D. (2001). Does repetition improve consistency? *Experimental Economics 4*(1), 5–54.

Hey, J. D. and C. Orme (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica 62*(6), 1291–1326.

Hill, R. V. and A. Viceisza (2012). A field experiment on the impact of weather shocks and insurance on risky investment. *Experimental Economics 15*(2), 341–371.

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

Humphrey, S. J. and A. Verschoor (2004). Decision-making under risk among small farmers in East Uganda. *Journal of African Economies 13*(1), 44–101.

IBLI PIU (2012). Index-Based Livestock Insurance (IBLI) project implementation report 2005-2012. IBLI Project Implementation Unit (PIU), Ulaanbaatar.

IFRC and MRCS (2010). Mongolia information bulletin no 1: cold waves. International Federation of Red Cross and Red Crescent Societies (IFRC) and Mongolian Red Cross Society (MRCS), Ulanbaatar.

Jacobson, S. and R. Petrie (2009). Learning from mistakes: What do inconsistent choices over risk tell us? *Journal of Risk and Uncertainty 38*(2), 143–158.

Jaeger, D. A., T. Dohmen, A. Falk, D. Huffman, U. Sunde, and H. Bonin (2010). Direct evidence on risk attitudes and migration. *The Review of Economics and Statistics 92*(3), 684–689.

Janzen, S. A. and M. R. Carter (2013). After the drought: The impact of microinsurance on consumption smoothing and asset protection. NBER Working Paper 19702, National Bureau of Economic Research, Cambridge, MA.

Jasso, G. and B. Wegener (1997). Methods for empirical justice analysis: Part 1. Framework, models, and quantities. *Social Justice Research 10*(4), 393–430.

Kachelmeier, S. J. and M. Shehata (1992). Examining risk preferences under high monetary incentives: Experimental evidence from the people's republic of china. *The American Economic Review 82*(5), 1120–1141.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), 263–291.

Karlan, D., R. Osei, and I. Osei (2014). Agricultural decisions after relaxing credit and risk constraints. *Quarterly Journal of Economics 129*(2), 597–652.

Kingdon, G. G. and J. Knight (2007). Community, comparisons and subjective well-being in a divided society. *Journal of Economic Behavior & Organization 64*(1), 69–90.

Klasen, S. (2000). Measuring poverty and deprivation in South Africa. *Review of Income and Wealth 46*(1), 33–58.

Knight, J., L. Song, and R. Gunatilaka (2009). Subjective well-being and its determinants in rural China. *China Economic Review 20*(4), 635–649.

Kolenikov, S. and G. Angeles (2004). *The use of discrete data in PCA: theory, simulations, and applications to socioeconomic indices*. Chapel Hill, North Carolina, US: Carolina Population Center, University of North Carolina.

Kolenikov, S. and G. Angeles (2009). Socioeconomic status measurement with discrete proxy variables: Is principal component analysis a reliable answer? *Review of Income and Wealth 55*(1), 128–165.

Kuhn, P., P. Kooreman, A. Soetevent, and A. Kapteyn (2011). The effects of lottery prizes on winners and their neighbors: Evidence from the Dutch postcode lottery. *The American Economic Review 101*(5), 2226–2247.

Kuznets, S. (1955). Economic growth and income inequality. *The American Economic Review 45*(1), 1–28.

Law of Mongolia (2014). Index-based livestock insurance law. State Information Publication Booklet 26/839, Ulanbaatar.

Lerman, R. I. and S. Yitzhaki (1985). Income inequality effects by income source: A new approach and applications to the United States. *The Review of Economics and Statistics 67*(1), 151–156.

Liebenehm, S. (2015). *New Insights from behavioral economics in developing countries*. Ph. D. thesis, Leibniz Universität Hannover.

Liebenehm, S. and H. Waibel (2014). Simultaneous estimation of risk and time preferences among small-scale cattle farmers in West Africa. *American Journal of Agricultural Economics 96*(5), 1420–1438.

Likert, R. (1974). A method of constructing an attitude scale. In G. M. Maranell (Ed.), *Scaling: A Sourcebook for Behavioural Scientists*, pp. 233–243. Transaction Publishers, New Jersey.

Linssen, R., L. van Kempen, and G. Kraaykamp (2011). Subjective well-being in rural india: The curse of conspicuous consumption. *Social Indicators Research 101*(1), 57–72.

Liu, E. M. (2013). Time to change what to sow: Risk preferences and technology adoption decisions of cotton farmers in China. *The Review of Economics and Statistics 95*(4), 1386–1403.

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

Loomes, G. and G. Pogrebna (2014). Measuring individual risk attitudes when preferences are imprecise. *The Economic Journal 124*(576), 569–593.

Luttmer, E. (2005). Neighbors as negatives: Relative earnings and well-being. *Quarterly Journal of Economics 120*(3), 963–1002.

Macours, K. (2011). Increasing inequality and civil conflict in Nepal. *Oxford Economic Papers 63*(1), 1–26.

Maddala, G. (1986). *Limited-dependent and qualitative variables in econometrics*. Number 3. Cambridge, UK: Cambridge University Press.

Maertens, A., A. V. Chari, and D. R. Just (2014). Why farmers sometimes love risks: Evidence from India. *Economic Development and Cultural Change 62*(2), 239–274.

Mahul, O. and J. Skees (2007). Managing agricultural risk at the country level: The case of Index-Based Livestock Insurance in Mongolia. World Bank Policy Research Paper 4325, World Bank, Washington D.C.

Mc Bride, M. (2010). Money, happiness, and aspirations: An experimental study. *Journal of Economic Behavior & Organization 74*(3), 262–276.

McInish, T. H., S. N. Ramaswami, and R. K. Srivastava (1993). Do more risk-averse investors have lower net worth and income? *Financial Review 28*(1), 91–106.

Menkhoff, L. and S. Sakha (2014a). Multiple-item risk measures. Kiel Working Paper 1980, Kiel Institute for the World Economy.

Menkhoff, L. and S. Sakha (2014b). Risk aversion over time: Experimental evidence in rural Thailand. Unpublished Manuscript.

Miranda, M. J. and K. Farrin (2012). Index insurance for developing countries. *Applied Economic Perspectives and Policy 34*(3), 391–427.

Mobarak, A. M. and M. Rosenzweig (2012). Selling formal insurance to the informally insured. Working Paper, Number 97, Department of Economics, Yale University.

Mobarak, A. M. and M. Rosenzweig (2014). Risk, insurance and wages in general equilibrium. NBER Working Paper 19811, National Bureau of Economic Research, Cambridge, MA.

Mobarak, A. M. and M. R. Rosenzweig (2013). Informal risk sharing, index insurance, and risk taking in developing countries. *The American Economic Review 103*(3), 375–380.

Moghaddam, F. M. (2005). The staircase to terrorism: A psychological exploration. *American Psychologist 60*(2), 161.

Mongolian Statistical Information Service (2015). Statistical database by sector. `http://www.1212.mn`. Accessed: 2015.08.15.

Moulton, B. (1990). An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *The Review of Economics and Statistics*, 334–338.

Murphy, D. J. (2011). *Going on otor: disaster, mobility, and the political ecology of vulnerability in Uguumur, Mongolia*. Ph. D. thesis, Department of Anthropology, University of Kentucky.

Nielsen, T., A. Keil, and M. Zeller (2013). Assessing farmers' risk preferences and their determinants in a marginal upland area of Vietnam: A comparison of multiple elicitation techniques. *Agricultural Economics 44*(3), 255–273.

NSO (2011). *Statistical Yearbook 2010*. National Statistical Office Mongolia, Ulaanbaatar.

NSO (2013). *Statistical Yearbook 2012*. National Statistical Office Mongolia, Ulaanbaatar.

Permanyer, I. (2014). Assessing individuals' deprivation in a multidimensional framework. *Journal of Development Economics 109*(4), 1–16.

Powdthavee, N. (2009). How important is rank to individual perception of economic standing? a within-community analysis. *The Journal of Economic Inequality 7*(3), 225–248.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica 32*(1), 122–136.

Runciman, W. (1966). *Relative deprivation & social justice: a study of attitudes to social inequality in twentieth-century England*. London, UK: Routledge Kegan Paul.

Saha, A. (1993). Expo-power utility: A flexible form for absolute and relative risk aversion. *American Journal of Agricultural Economics 75*(4), 905–913.

Scheibehenne, B., R. Greifeneder, and P. M. Todd (2010). Can there ever be too many options? A meta-analytic review of choice overload. *Journal of Consumer Research 37*(3), 409–425.

Schoemaker, P. J. (1982). The expected utility model: Its variants, purposes, evidence and limitations. *Journal of Economic Literature 20*(2), 529–563.

Shah, A. K., S. Mullainathan, and E. Shafir (2012). Some consequences of having too little. *Science 338*(6107), 682–685.

Shaw, K. L. (1996). An empirical analysis of risk aversion and income growth. *Journal of Labor Economics 14*(4), 626–653.

Skees, J. and A. Enkh-Amgalan (2002). Examining the feasibility of livestock insurance in Mongolia. World Bank Policy Research Paper 2886, World Bank, Washington D.C.

Skees, J. R. and B. J. Barnett (2006). Enhancing microfinance using index-based risk-transfer products. *Agricultural Finance Review 66*(2), 235–250.

Stark, O. (1984). Rural-to-urban migration in LDCs: a relative deprivation approach. *Economic Development and Cultural Change 32*(3), 475–486.

Stark, O. and E. Taylor (1991). Migration incentives, migration types: The role of relative deprivation. *Economic Journal 101*(408), 1163–1178.

Starmer, C. and R. Sugden (1991). Does the random-lottery incentive system elicit true preferences? An experimental investigation. *The American Economic Review 81*(4), 971–978.

Sternberg, T. (2010). Unravelling Mongolia's extreme winter disaster of 2010. *Nomadic Peoples 14*(1), 72–86.

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

UNDP and NEMA (2010). Dzud National Report 2009-2010. United Nations Development Program (UNDP) and National Emergency Management Agency (NEMA), Ulaanbaatar.

Van Praag, B., A. Hagenaars, and W. van Eck (1983). The influence of classification and observation errors on the measurement of income inequality. *Econometrica: Journal of the Econometric Society 51*(4), 1093–1108.

Vieider, F. M., M. Lefebvre, R. Bouchouicha, T. Chmura, R. Hakimov, M. Krawczyk, and P. Martinsson (2015). Common components of risk and uncertainty attitudes across contexts and domains: Evidence from 30 countries. *Journal of the European Economic Association 13*(3), 421–452.

Von Neumann, J. and O. Morgenstern (1953). *Theory of games and economic behavior*. Oxford University Press.

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

Wilcox, N. T. (2008). Stochastic models for binary discrete choice under risk: A critical primer and econometric comparison. In J. Cox and G. Harrison (Eds.), *Risk Aversion in Experiments*, Volume 12, Chapter 7, pp. 197–292. Emerald Group Publishing Limited.

Wooldridge, J. M. (2007). Inverse probability weighted estimation for general missing data problems. *Journal of Econometrics 141*(2), 1281–1301.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT Press.

World Bank (2013). Implementing agriculture for development: World Bank Group Agriculture Action Plan (2013-2015). World Bank, Washington D.C.

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

Yitzhaki, S. (1979). Relative deprivation and the Gini coefficient. *Quarterly Journal of Economics 93*(2), 321–324.

Zimmerman, F. J. and M. R. Carter (2003). Asset smoothing, consumption smoothing and the reproduction of inequality under risk and subsistence constraints. *Journal of Development Economics 71*(2), 233–260.

Statement of Contributions

Chapter 2 is co-authored with Ghassan Baliki. The study conception and the writing were joint work. The empirical analysis was led by Veronika.

Chapter 3 is co-authored with Karlijn Morsink. Financial support for the experiment was acquired and the field-work was coordinated by Karlijn. The experimental design, the study conception, the empirical analysis and writing was led by Veronika.

Chapter 4 is co-authored with Kati Krähnert. Financial and administrative support for the data collection was procured by Kati. The study conception and writing were joint work. The empirical analysis was led by Veronika.

EIDESSTATTLICHE ERKLÄRUNG

Ich versichere, dass ich die von mir vorgelegte Dissertation selbstständig und ohne unerlaubte Hilfe angefertigt habe und mich keiner anderen als der in ihr angegebenen Hilfsmittel bedient zu haben. Insbesondere sind sämtliche Zitate aus anderen Quellen als solche gekennzeichnet und mit Quellenangaben versehen.

Berlin, 12. September 2015

Acknowledgements

This thesis would not have been possible without the generous help and input of many people. I am deeply thankful for the valuable discussions throughout my doctoral research, which were crucial in shaping and refining my thinking and writing.

Most importantly, I am indebted to my supervisors, Professor Susan Steiner and Professor Nathan Fiala, who supported me greatly and gave me the freedom to follow my research interests and to grab my opportunities. My co-authors Kati Krähnert, Karlijn Morsink and Ghassan Baliki, I would like to thank for the inspiring and fruitful collaboration. It has truly been a privilege to work with them.

Many others have provided valuable comments and support. I am grateful to my colleagues at the D&S department of DIW Berlin, my fellow students in the DIW Graduate Center and the DENeB. In particular, I would like to thank Friederike Lenel who helped me to grow in numerous debates and critical reflections. Furthermore, I would like to thank Christopher Boyce, Andrew Clark, Conchita D'Ambrosio, Glenn Harrison, Kriztina Kis-Katos, Andreas Landmann, Adam Lederer, Wei Liu, Mushfiq Mobarak, Dan Osgood, Jerry Skees, and Ferdinand Vieider for their valuable input.

Many thanks for receiving financial support at different stages for this research from DIW Berlin, the German member organization of the International Institute for Applied Systems Analysis (IIASA) Laxenburg, and Gottfried Wilhelm Leibniz Universität Hannover.

I would like to thank Christoph for encouraging me to step on new paths and for believing in me and my aspirations. I thank Nuria for distracting me in the most wonderful way.

Last but not least, I would like to thank more than 5,000 individuals in Kyrgyzstan and western Mongolia, who answered patiently the survey questions, and in the Tigray region of Ethiopia, who participated in the artefactual field experiment.