GOTTFRIED WILHELM LEIBNIZ UNIVERSITÄT
HANNOVER

Essays on Risk Preferences and Peer Effects in Household Decision-Making: Experimental Evidence

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

Doktorin der Wirtschaftswissenschaften
– Doktor rerum politicarum–

genehmigte Dissertation

von

Master of Science Sahra Sakha

geboren am 06.05.1988 in Essen

Januar 2016

Referent: Prof. Dr. Lukas Menkhoff, Deutsches Institut für Wirtschaftsforschung (DIW)

Koreferent: Prof. Dr. Stephan Klasen, Georg-August-Universität Göttingen

Tag der Promotion: 15.01.2016

Abstract

Behavioral economics has improved the theory of human behavior and provided new psychological underpinnings for economics analysis. For example, there is nothing in traditional theory that specifies that people should care about fairness, should weight risky outcomes in a linear fashion or that they must discount their future exponentially at a constant rate. Using these anomalies inspired to create alternative theories that could generalise existing theoretical models and suggested new policy options.

Classical economic theory is based on a number of strong and testable assumptions. Three commonly made assumptions are:

I Risk preferences are constant over time and in particular are unaffected by personal circumstances (state-independence).

II Risk preferences do not vary according to the context in which they are revealed (context-invariance).

III Consumption choices are a function only of budget, price and static personal preferences and in particular unaffected by an individual's social context (no peer effects).

The empirical results I present in this thesis call the appropriateness of all three of these assumptions into question. Risk preferences are neither state-independent nor context invariant. Peer effects on individual behaviour appear both deep and pervasive. The picture that emerges is one of individuals embedded in a network of social relationships that shape their preferences, incentives and constraints. Hence, research on the social context of household economic decision making is useful to better predict behavior and to make meaningful policy impact assessments.

In what follows, I review the literature on risk preferences and peer effects to place the chapters of the thesis into context. The first two chapters of my thesis discuss topics in decision-making under risk and uncertainty. They contest the first two assumptions mentioned above — that preferences are state-independent and context-invariant. The last chapter provides an insight into decision-making considering peer effects.

There are three clear conclusions from my research. Chapter 1 shows that risk preferences are not state-independent but rather alter over time with respect to micro-economic shocks, changes in well-being, and the macroeconomic environment. Chapter 2 shows that risk preferences are not context-invariant: different risk elicitation tasks lead to inconsistent measures of risk aversion, making it difficult to predict risky behaviour based on a single measure. In addition, we provide a new risk preference measure to extrapolate real-world behavior in many dimensions in a more robust and consistent way. Chapter 3 shows that peer effects exist in consumption decisions. All chapters use lab-in-the-field experiments to test predictions.

Keywords: Risk preferences, risky behavior, peer effects

Kurzfassung

Die Verhaltensökonomik hat die Theorie menschlichen Verhaltens bereichert und mit neuen psychologische Aspekten eine weitere Basis für Wirtschaftsanalysen geschaffen. Damit hat sie die Ökonomie vorangetrieben. Beispielsweise gibt es nichts in traditionellen Theorien, die zeigt, dass Menschen zu Gerechtigkeit und Fairness neigen, dass sie risikoreiche Folgen nicht-linear gewichten oder dass sie ihre Zukunft exponentiell mit einer konstanten Rate diskontieren. Diese neuen Erkenntnisse haben die Vorhersagefähigkeit der Wirtschaftstheorien verbessert und mögliche neue Strategien aufgezeigt.

Starke und prüfbare Thesen sind die Grundlage der klassischen Wirtschaftstheorie. Drei häufig aufgestellte Thesen sind:

I Die Risikobereitschaft bleibt über die Zeit konstant und insbesondere von persönlichen Umständen unbeeinflusst (Zustands-Unabhängigkeit).

II Die Riskikobereitschaft bleibt in dem Kontext, in dem sie entstanden ist, unverüdert (Kontext-Beständigkeit).

III Konsumentscheidungen sind nur eine Funktion des Budgets, des Preises und der persönlichen Vorlieben. Sie bleiben vom sozialen Umfeld unbeeinflusst (keine Peer-Effekte).

Die Richtigkeit aller drei genannten Thesen wird mit den empirischen Ergebnissen dieser Arbeit in Frage gestellt; die Risikobereitschaft ist weder unabhängig von der Zeit noch Kontext beständig. Peer-Effekte scheinen eine tiefgehende und überzeugende Wirkung auf individuelle Verhaltensweisen zu haben. Es zeigt sich, dass Vorlieben, Anreize und Einschränkungen jedes Einzelnen in einem Netzwerk sozialer Beziehungen geformt werden. Demzufolge ist die Erforschung des sozialen Umfeldes und dessen Einfluss auf die wirtschaftlichen Entscheidungen eines Haushalts sinnvoll, um Verhaltensweisen besser vorauszusagen und um die

Auswirkung von politisch bedeutsamen Strategien einzuschätzen.

Die Literaturauswertung in Bezug auf Risikobereitschaft und Peer-Effekte wird die folgenden Kapitel dieser Arbeit in Zusammenhang stellen. Die ersten zwei Kapitel setzen sich mit Themen der Entscheidungsfindung unter Risiken und Ungewissheit auseinander. Thesen (1) und (2) obenerwähnt werden in Frage gestellt. Das letzte Kapitel stellt Erkenntnisse zur Entscheidungsfindung unter Berücksichtigung des Peer-Effekts vor.

Drei klare Folgerungen werden in dieser Arbeit getroffen. Kapitel 1 zeigt, dass die Risikobereitschaft nicht unabhängig vom Zustand ist, jedoch sich im Laufe der Zeit hinsichtlich mikroökonomischen Schocks, Veränderungen im Wohlbefinden und des makroökonomischen Umfelds verändert. Kapitel 2 zeigt, dass eine Risikobereitschaft nicht kontextbeständig ist: verschiedene Maße der Risikoerhebung führen zu widersprüchlichen Ergebnissen der Risikoaversion. Diese erschweren eine Voraussage zum Risikoverhalten, basierend auf einer einzigen Messgröße. Zudem wird eine neue, robustere und beständigere Messung der Risikobereitschaft zur Fortschreibung von Verhaltensweisen vorgestellt. Kapitel 3 zeigt, dass Peer-Effekte zu Konsumentscheidungen existieren. In allen Kapiteln werden Feldexperimente zur Prüfung der Thesen genutzt.

Schlagwörter: Risikoeinstellung, Risikoverhalten, Peer-Effekte

Acknowledgements

First and foremost I would like to thank my advisor Lukas Menkhoff. It has been an honor to be his Ph.D. student. I thank him for his tremendous academic support, which made my Ph.D. experience productive and stimulating. He provided me with the opportunity to conduct my peer experiment in rural Thailand and participate in various seminars and conferences. He has always been available to guide me over three years of thesis development. During the most challenging times when writing this thesis, he gave me the encouragement and freedom I needed to move on.

I would also like to extend my thanks to my committee members, Stephan Klasen, Susan Steiner, and Attakrit Leckcivilize for serving as my committee members and providing me with invaluable academic inputs for my research papers throughout my Ph.D. The members of the Research Training Group (RTG) have contributed immensely to my personal and professional time in Hannover and Goettingen. The group has been a source of friendships as well as good advice and collaboration. I am also grateful to other faculty members in Hannover. I am especially grateful to: Karoline David (for the support and encouragement whenever I was in need), Kai Gehring (for proofreading my thesis and giving helpful comments on my papers) and Tabea Herrmann (for spreading happiness on my scientifically dark days). I would like to acknowledge my co-author Antonia Grohmann. We worked together on the peer effects experiment, and I greatly appreciate her enthusiasm, intensity, and inspiring discussions over countless coffee breaks that made the project possible.

I am also indebted to my other colleagues and friends who contributed immensely to my personal and professional advancement over the years and in the final stages of my thesis, in particular Shriya Bubna, Francesca Cristaldi, Andrea Deghi, Cirus Iniesta, Juan David Laguna, and Ben Waltmann.

I gratefully acknowledge the funding sources that made my Ph.D. work possible. I was funded by the Deutsche Forschungsgemeinschaft (DFG 756) for my first year, the Friedrich-Ebert-Stiftung, for years two and three and the RTG 1723.

My time in Hannover was made enjoyable in large part due to the many friends and groups that became a part of my life. I am grateful for time spent with colleagues and friends and for my memorable trips in rural Thailand where I walked through rice paddies to get to my respondents or drove 70 km to lay my hands on a particular candy for the tasty treat. It was an unforgettable experience – from the beginning until the end.

I would like to thank Manuel Rojas Soto whose faithful support, devotion, and motivation during the final stages of this Ph.D. is so appreciated. Lastly, I would like to thank my family for all their love and encouragement. Their support has been unconditional all these years; they have given up many things for me so that I could be where I am now; they have cherished with me every great moment and supported me whenever I needed it. This thesis is dedicated to you.

Introduction

In the expected utility (EU) framework and its variants including prospect theory (Kahneman and Tversky, 1979), risk attitude is a descriptive label for the shape of the utility function presumed to underlie a person's choices. A person's risk attitude describes the shape of his or her utility function (derived from a series of risky choices) for the outcomes in question. A commonly used metric of risk attitude is defined as $-u''(x)/u'(x)$, where u' and u'' denote the first and second derivatives of the utility function u, respectively (Arrow, 1974; Pratt, 1964). The terms 'risk averse' and 'risk seeking' within the EU framework technically refer only to the curvature of the utility function. Yet, those who coined the term risk aversion had in mind the psychological interpretation that someone who prefers the expected value of a gamble over playing the gamble does not like to take risks – an interpretation which has a lot of currency among researchers and the general public. Risk attitude, a person's standing on the continuum from risk aversion to risk seeking, is commonly considered to be a stable characteristic of individuality, a kind of personality trait. For instance, Stigler and Becker (1977) argue that preferences in general and risk preferences in particular do not "change capriciously" (p. 76).

Recently economic researchers have challenged the interpretation of risk attitudes in the EU-sense as a personality trait, arguing that preferences are contingent on the context in which they are expressed. Levitt and List (2007) argue that real-world behavior is not just influenced by monetary calculations, but also other factors such as the presence of moral and ethical considerations; the nature and extent of scrutiny of one's actions by others; and the context in which the decision is embedded.

This has wide ranging real-world implications. Consider a situation in which an individual has to decide over a pension plan when entering a job market at a young age. Hence, for the decision-maker it is essential that the decision made today fits her preferences when the payoff materializes in the future. Therefore, either the preferences need to be stable between the time of decision and the payoff or the decision maker has to anticipate correctly a change in the preferences over time when making the decision. The assumption of temporal stability of risk preferences is, therefore, crucial and whether this assumption is reasonable for real-world preferences is ultimately an empirical question.

Empirical evidence suggests that individuals' risk preferences are not state independent and can be altered by, for instance, negative shocks such as hyper-inflation (Malmendier and Nagel, 2011), financial crises (Guiso et al., 2014), conflicts or violent trauma (Voors et al., 2012; Callen et al., 2014), and natural disasters (Cameron and Shah, 2015). There is little consensus, however, as to whether negative shocks induce people to become more risk averse or risk tolerant. The evidence from existing studies that exploit a variety of negative shocks is mixed and mostly based on cross-sectional data.

Chapter 1 contributes to a nascent body of literature on the stability of risk preferences. However, the existing literature has largely focused either on the impact of aggregate shocks (i.e. natural disasters, Cameron and Shah, 2015) or idiosyncratic shocks (i.e. personal financial situation, Andersen et al., 2008b). Our study makes several contributions over previous research. First, we not only look at the impact of changes on risk aversion at the idiosyncratic level by looking at socio-economic characteristics, but also disentangle the impact of micro and macroeconomic influences on changes in risk aversion. Second, as a measure of risk aversion, we employ a repeated incentivized certainty equivalent experiment. We can thus use panel data to identify the influence of past behavior on current behavior and control for unobserved individual fixed and time-varying characteristics. Third, even after five years we are able to maintain a relatively large sample which is drawn from a representative subject pool. Finally, we are able to control for attrition and selection bias in conjunction with the fixed effects model. As a result, the regression outcomes can reasonably be given a causal interpretation.

The experiment is administered as part of a large household survey from the ongoing panel project "Impact of shocks on the vulnerability to poverty in Southeast Asia" and was conducted in the largest province of Thailand, Ubon Ratchathani. We choose rural Thailand because it is widely recognized that risk and uncertainty have a stronger impact on household decision making in developing countries because insurance provision is still limited. We conducted the certainty equivalent experiment in 2008, 2010 and 2013 involving 384 participants across all three waves. The link to the large household survey adds not only wider socio-economic information but also expectations (i.e. optimisms with regard to the future outlook) and personal shocks of the households (e.g. death of household member, collapse of business, crops pests etc.). An indicator for the general macroeconomic environment is constructed using time dummies.

The analyses reveal that individual risk aversion changes significantly over time. Correlations between risk measures are weak and indicate an overall decrease in risk aversion over time. Investigating determinants for individual-level preference stability using a fixed effects model, we find that variation in risk attitudes is not significantly correlated with socio-economic characteristics. Instead, we find that individuals' perceived wellbeing and the number of shocks they experience have a negative effect on risk aversion. A more pessimistic outlook of the future is associated with an increase in the level of risk aversion. Examining the influence of shocks by type in a separate analysis highlights the importance of economic shocks on preference stability (agricultural shocks are only marginally significant). The macroeconomic environment instead seems to have limited influence. Furthermore, we find that the impact of the number of economic shocks on changes on risk preferences is even stronger for low income and low insured households. The chapter concludes that exposure to shocks has a large impact on people's risk-taking behavior. The results on insurance presented above point to one potential policy solution.

Another assumption that is deeply rooted in the EU framework, yet at odds with empirical observations, is the notion of context-invariant risk preferences. Standard models in many fields of economics generally use one canonical model for decisions under uncertainty, in which individuals (or households) have a single, concave utility function over wealth, which gives rise to context-invariant risk preferences. Guided by this assumption, standard practice is to use one estimate of the risk aversion parameter to calibrate the model. However, at the other end of the spectrum, there is a large literature in psychology and behavioral economics arguing that context is king. Empirically, there is little, if any, commonality in how the same individual makes decisions across different contexts and different tasks. Where does reality lie relative to these two extremes?

Different methods of measuring people's utility have been shown to result in different classifications (Slovic, 1964). The certainty-equivalent method, for example, results in utility function with more extreme risk attitudes (more risk aversion for gains, more risk seeking for losses) than the probability equivalence method. Furthermore, individuals do not appear to be consistently risk seeking or risk averse across different domains and situations even when using the same assessment method (Hey at al. 2009). From a practical viewpoint, this questions the usefulness of the methods to control for previously determined risk aversion in experimental games and its ability in predicting real-world behavior. So how can researchers deal with this inconsistency? In many cases it is difficult to decide ex ante which specific risk measure is most appropriate or whether one risk measure is superior to another in eliciting risk attitude. This is the starting point of our research.

Chapter 2 provides new empirical evidence. Our approach differs from this literature by comparing not only the within-subject variability of responses across seven established risk elicitation methods but also relate these risk tasks to five dimensions of risky behavior (i.e. gambling, self-employment, investment, insurance, and health) using the in-depth household survey "Impact of shocks on the vulnerability to poverty". We aim to investigate which risk measure or combination of risk measures may be most broadly applicable to different kinds of risky behavior. The benchmark for deciding on the effectiveness of a risk measure is its power in explaining risk-related behavior while controlling for individual socio-demographic characteristics.

While all risk measures have some power in explaining risky behavior in our sample, there is enormous heterogeneity, and correlations among risk elicitation tasks seem low. In terms of the explanatory power of the risk tasks, some measures perform better than others. We also find domain specificity: risk tasks which were framed in a financial context were much better in explaining financial behavior (i.e. investment). Hence, we can reject the idea of context-invariant preferences. The question that naturally arises from this finding is, if risk tasks have varying degrees of predictability on real-world behavior, is there a general risk measure more broadly applicable to many dimensions of risky behavior? The second contribution of our research is to create a new risk measure that robustly and consistently explains a wider range of risky behavior. We create a multiple-item risk measure by averaging across the seven risk elicitation measures. This improves the predictive power in explaining behavior substantially. We further find that analyzing the effect of combining risk items with different framings contributes to a more reliable and predictive multiple-item risk measure. This is in contrast to the combination of items from different domains or averaging across repeated answers, which seems less helpful in improving predictive power. Evidence on whether and how elicited estimates of preferences over uncertainty correlate with real-world behavior is important in order to advance our knowledge on the external validity of such measures.

The standard economic approach to the analysis of choice under risk emphasizes the role of individual risk preferences. In deciding how much to invest in a risky asset, for instance, individuals weigh up the costs and benefits referring to these preferences. By contrast, in many important settings individuals do not take choices in isolation and the social settings within which choices are made influence behavior. For example, individual choices in isolation may be swayed by the opinions and decision of others. This feeling is familiar to anyone who has ever been transfixed with infatuated desire for a product even though the price seemed too high, the budget was tight, and the item was not desperately needed. Yet one is convinced otherwise because another person bought the same product or the product was recommended by a member of one's peer group.

Consumption is arguably a social experience, and the position of other people with re-spect to our own consumption often matters to us. However, relatively little is known about the impact of group behavior on the individual in terms of consumption choices. The essential problem in studying the causal impact of peer effects is that, regardless whether the researcher is interested in how individual behavior is affected by group characteristics (termed exogenous or contextual effects) or group behavior (termed endogenous effects), data are rarely available in which the relevant groups or their associated traits are exogenously assigned. While this criticism applies to any empirical study when we examine how individual traits are associated with individual outcomes, the problem is particularly vexing in the study of peer effects. The conceptual problems are numerous and well elucidated in the literature. Manski (1993, 2000) points out numerous pitfalls whereby a researcher may erroneously infer the presence of peer effects, when in fact the estimates may only be indicative of the respondent and her associated group sharing a common environment. Hence, identifying peer effects from observational data is highly difficult (Angrist and Lang, 2004).

Chapter 3 proposes a novel strategy for the identification and estimation of peer effects on consumption and, particularly, of their endogenous component using an experimental design in a fully controlled setting where no possible confounding factor can hinder identification. To the best of our knowledge, no experiment of peer effects in consumption decisions has been conducted so far. The design of our experiment is straightforward and was implemented in Thailand: we test consumption choices by simply offering respondents the choice between a combination of sweet and salty snacks, i.e. the tasty treat, and money. The amount of money offered increases in every round by ten Baht whereas the tasty treat stays the same. In the control group, respondents have to make their consumption choices on their own, separated from the rest of the respondents. In the treatment group, each respondent still makes his/her own decision, but all respondents play whilst observing each other. We randomly assigned the villages to their respective treatments. Hence, the only difference in outcome can be attributed to peer observation.

We are able to control for personal and local confounding factors because our experimental results are complemented with a large household survey. We overcome the problem of correlated effects because randomization of our sample in observing and non-observing groups on the village level combined with detailed information about village and household characteristics circumvents the problem of contextual effects. Thus, our research design enables us to directly compare outcomes for those groups that performed the experiment with and without peer observation. We use the model of endogenous peer effects with leave-out mean in which individuals' average consumption is regressed on the mean of the group average – excluding the individual himself (Angrist, 2014).

First, we find that the standard deviations of those groups that observe each other are lower than for those groups that do not observe each other. Hence, individual's make more converging decisions in terms of consumption choices -– which is in line with the predictions of our model. At the same time, we show that individual choices are higher when the leave-out group mean is higher. Most importantly, we only observe this when the experiment is performed with peer observations. Hence, we provide clear evidence of peer effects and conclude that peer observation leads to conformity. We further study the effect of familiarity with the product and find that peer observation can counteract the effect of a lack of knowledge of a product. Looking into treatment heterogeneities, we find that individuals with high cognitive ability compared to their group, are less likely to choose the tasty treat, while the same is not found for high-income and overconfident respondents. We find that the magnitudes of peer effects of individuals in smaller villages and more remote areas are more pronounced, rather than villages with more inhabitants who are also closer to the city.

Why is the study of peer effects consumption important? From a welfare point of view, one may be interested in measuring and understanding the type of distortions (if any) induced by the presence of peer effects. Depending on the mechanism underlying peer effects, distortions may be induced by the conspicuous consumption and /or by the keeping-up with-the-Joneses channel. In the first case, budget shares would be distorted, i.e. status-seeking behavior might impact the share of conspicuous goods in the consumption bundle. Since conspicuous goods are typically luxuries, consumption peer effects might have noticeable welfare consequences in the form of excess "wasteful" consumption. This may induce under-saving (or over-borrowing) in the attempt to keep up with the Joneses. Thus, the analyses of peer effects help to identify why households fail to save optimally.

My thesis contributes to behavioral economics and makes significant contributions to wards understanding the deeper mechanisms behind household's financial decision-making. The overall result is that behavioral aspects matter. Risk preferences are not stable as assumed, neither over changes in the states of nature, nor over different contexts. Furthermore, we find that decision-making of the individual critically depends upon the decisions and judgements of other group members. In each case, the incorporation of behavioral factors complements existing economic theories by making them more practical, tractable and thus more general.

Contents

Chapter 1

Determinants of Risk Aversion over Time: Experimental Evidence

[1] This chapter is based on the article by Menkhoff and Sakha (2015). Thanks go to participants at the European Economic Association Conference in Toulouse, the Spring Meeting of Young Economists in Vienna, and the 9th Nordic Conference on Behavioral and Experimental Economics for helpful comments and discussions. We thank Alessandro Bucciol, Oliver Gloede, Bernd Hardeweg, Veronika-Bertram Huemmer, Astri Muren, Stephan Klasen, Friederike Lenel and Susan Steiner. Financial support by the German Research Foundation (DFG projects: RTG 1723 and ME 1070) is gratefully acknowledged.

1.1 Introduction

In a world of rational decision makers, and assuming no changes in the state of the world, risk preferences are supposed to be stable over time. More precisely, when asked the same question at separated moments in time, we expect most subjects to provide the same answer. We would like to think that risk preferences fall in the category of fundamental underlying preferences that define our personality and identity. Stability of risk preferences is also an assumption made when we elicit preferences for subjects. In areas such as financial investment, risk insurance, and health decision, clients are often asked about their risk preferences. The recommendations the client receives hinge on the assumption that risk preferences are stable.

In contrast to the importance of stable risk preferences over time, we know relatively little about the degree of stability and potential determinants of its change. Some studies indicate that there are significant changes when repeatedly measuring individual risk aversion within short time periods (Hey, 2001, Guiso et al., 2014). On the contrary, Sahm (2012) and Chiappori and Paiella (2011) using panel data found no changes in risk attitude over time.

With regard to the determinants of risk aversion, it seems that an improvement or worsening of the idiosyncratic and aggregate environment has an impact on the evolution of risk attitude. On the idiosyncratic level, Andersen et al. (2008b) find that people become more risk tolerant when their personal financial situation ameliorates. In addition, Malmedier and Nagel (2011) find that individuals' experiences of macro-economic outcomes such as the Great Depression have long-term effects on their risk attitudes. On the aggregate level, Guiso et al. (2014) find similar results, however, in the other direction. Using a repeated survey of a large sample of clients of an Italian bank, they find that the deterioration of the overall macroeconomic condition with respect to the recent financial crisis leads to an increase in risk aversion. Further, they find that the results of changes in the macroeconomic environment are highly correlated with psychological factors such as fear that seems to drive changes in risk aversion.

There is also an emerging literature that investigates changes of risk preferences by extreme events such as natural disasters or civil wars. Main results of the impact of natural disasters and civil wars are mixed, some report it makes people more risk averse (Cameron and Shah, 2015; Callen et al., 2014) while others report the opposite (Bchir and Willinger, 2013). Bchir and Willinger (2013), for instance, consider the volcanic threat in Peru. They find that poor households are more risk seeking and more impatient in exposed areas than in unexposed areas, but not high income households. The main drawback of this literature, however, is that it is mainly based on cross-sectional data with no available data on preferences before and after the event.

A better understanding of the existing link between shocks and individuals' risk preferences might prove helpful to better target prevention, protection and emergency policies. Indeed, self-protection, emergency preparedness and self-insurance behavior are key issues related to individuals' risk preferences. Furthermore, from a theoretical point of view, natural hazards can be likened to background risk. Such types of risks are difficult to insure and hard to mitigate. Yet there is no clear theoretical prediction about how individuals' preferences are affected by background risk. According to expected utility theory (EUT), the most likely prediction is risk-vulnerability, that is individuals behave as if they are more risk-averse with respect to a given foreground risk in the presence of a harmful background risk than without (Gollier and Pratt, 1996).

While such shocks have obvious economic effects, they will also have potentially overlooked, important, and long-lasting consequences if they noticeably affect risk preferences and risk behavior. The objective of this paper is to assess, thanks to experimental and survey data, the precise impact of experiencing various microeconomic shocks on changes in risk preferences in Thailand.

We primarily analyze risk aversion changes between 2008 and 2013 by an incentivized experiment conducted in rural Thailand. The incentivized certainty equivalent task was used to measure risk aversion in Bruhin et al. (2010); Abdellaoui et al. (2011) and Dohmen et al. (2011). We complement the experiment with a non-incentivized risk survey items. The survey item is the general risk question from the German socio-economic panel study (SOEP) and used in Dohmen et al. (2011).

We choose rural Thailand because it is widely recognized that risk and uncertainty have a stronger impact on household decision making in developing than in developed countries. Rural households in developing countries face higher risks such as droughts, floods, pests or even health risks because insurance provision is still limited and state social safety nets are often not available. In such an environment, the poor have developed sophisticated risk coping mechanisms, involving activity diversification, mutual support networks and savings for precautionary purposes (Dercon and Christiaensen, 2011). However, the efficiency of such strategies has scope for improvement, reason for which it is believed that risk attitude variation over time should be more visible than in developed countries.

The purpose of our research is to achieve a better understanding of the determinants of changes in risk aversion. To evaluate such changes we have conducted a panel experiment of three waves over five years - involving 384 participants in all three waves. To the best of our knowledge, this data set is the first that enables us to explain risk aversion over a longer period of time, considering micro and macro effects together. This is feasible since experimental results can also be linked to a large household survey containing wider socio-economic information on the living conditions of the households. Even though this panel is not long enough to observe long-term developments, our focus on Thailand includes a period of unusually large macroeconomic fluctuations which allow some first inferences. This should reveal some variation of risk attitude in the cross-section and over time which could not be observed before. This data set allows us to use a panel approach in examining determinants of changes in risk aversion. The regression outcomes demonstrate about the impact of changes in socio-demographic variables, including well-being indicators and micro-economic shocks, on risk aversion after controlling for individual unobserved heterogeneity.

We find that individual risk aversion changes over time, which extends so far faint experimental evidence with respect to sample size and time period. Overall, it seems to decrease over time. This is evident for both experimental and non-experimental risk tasks. Determinants of such changes are of micro and macro origin. An improvement in the overall macroeconomic environment leads to a reduction in risk aversion. This coupled with microeconomic influences such as negative personal shocks and worsening well-being (or expectations thereof), increase risk aversion. These results hold in various robustness exercises. We also control for possible biases of endogenous sample selection that are inherent in virtually any panel experiments allowing subjects to drop out of the panel.

Regarding the repeated experiments on eliciting risk aversion, Hey and Orme (1994) are the first, to the best of our knowledge, to examine risk aversion over time. However, the repetition takes place the following days so that the issue is more on consistency than on time-variation. The most similar study is Andersen et al. (2008b) who elicit risk aversion from 253 Danish people and repeat the exercise with 97 of them once within the following 3 to 17 months. They find high variation in risk aversion over time, some of which can be weakly explained by changes in personal financial affairs but there is no link to macro factors. Another related paper is Chuang and Schechter (2015) who use experimental and survey measure to examine stability of risk, time, and social preferences. They find rather stable results for time and social preferences in terms of the survey items but not for the experimental results for risk preferences. However, they could not explain determinants of changes in risk aversion with changes in exogenous income shocks and it remains largely unanswered why they find significant changes in risk attitude (such results may be driven by noise, as noted by the authors).

Compared to this earlier work, we base our analysis on a repeated incentivized experiment and find more micro influences and a new macro influence. These findings seem to depend on the availability of a large sample, longer time-period and more volatile macro environment and thus do not contradict but extend earlier literature.

The remaining paper is structured as follows. Section 1.2 describes the data used in this research and investigates changes of risk attitude over time. Section 1.3 demonstrates first descriptive statistics. Section 1.4 shows how we control for attrition bias. Section 1.5 presents our identification strategy. Section 1.6 examines determinants of changes in risk attitude. Section 1.7 presents further robustness tests and Section 1.8 concludes.

1.2 Experiment and Experimental Results

This section describes the experiment revealing risk preferences. The experiment is embedded in a larger household survey. We start by describing the household survey in Section 1.2.1 and the certainty equivalent experiment in Section 1.2.3. We present the experimental outcome of all three waves in the cross-section in Section 1.2.3. Section 1.2.4 shows the experimental outcome over time.

1.2.1 Household Survey and Sampling Procedure

The experiment is administered as part of a large household survey from the project "Impact of shocks on the vulnerability to poverty: consequences for development of emerging Southeast Asian economies" which collects data from approximately 4,000 households in six provinces of Thailand and Vietnam starting in 2007.

The household selection process follows a three-stage stratified sampling procedure where provinces constitute strata and the primary sampling units are subdistricts. Within each province, we exclude the urban area around the provincial capital city and confine the sample to the remaining rural areas. Within each subdistrict, two villages are chosen at random, in which 10 households are randomly selected through the listing of registered citizens. Overall, the sampled households are representative for the rural areas in the considered province.

The household survey contains detailed information on socio-economic characteristics of the household and respondent including: household demographics, recurrent and durable expenditures, credit and savings, landholdings, agriculture, employment, health, education but also – and this is important for our studies - past experiences of microeconomic shocks and risks. It also includes materials concerning village characteristics such as the number of village institutions or infrastructure (i.e. irrigation system, number of social activities in the village, nursery, bank etc.), in and outward village migration, inhabitants, but also the number of shocks occurring in a village. The data set provides a representative sample of rural households in the Northeastern part of Thailand.

Contrary to the household survey, the experiment was conducted in Ubon Ratchathani, the largest of three provinces in Northeastern Thailand. The implementation process involved a pilot study with the purpose of reshaping the questionnaire for clarity. The final survey was implemented and respondents were interviewed face-to-face with a local enumerator. All enumerators were changed for each wave to exclude the possibility of enumerator fixed effects that could affect our results.

In the first wave in 2008 we cover 947 households - typically the household head is interviewed. In 2010, we re-interviewed 909 respondents while in the last series 2013, we still retained 851 participants. Overall, we have 384 respondents who participated in all three experiments in 2008, 2010 and 2013. For our analysis we employ this sub-sample.

1.2.2 Description of the Experiment and Risk Survey Items

The experiment is an incentivized certainty equivalent task to reveal risk preferences. An important concern in risk preferences elicitation through lottery games is the extent to which subjects understand the instructions. Even if we have selected a simple task, the experiment could be complex to understand, especially with field subjects in a developing country. As many of our participants have received little or no education, we have provided all experimenters with clear and visual instructions to make it easier for illiterate subjects to understand the consequences of any decisions they made in the game.

The structure of the experiment is given in Table A.2 (more details on the instruction are shown in Appendix A). It illustrates the basic payoff matrix presented to subjects. The first row shows that the lottery offers a 50-50% chance of receiving either 0 Thai Baht (THB) or 300 THB and alternatively a safe payoff of 0 THB. The expected value (EV) of this lottery is 150 THB constituting one day's labor wage. Therefore, it is rational to choose the lottery over the safe payoff. The second row, however, already offers 10 THB as safe payoff. This provides the opportunity for risk-averse individuals to opt for safe payoff over the lottery. The value of the safe payoff is increased in each row by 10 THB so that in the last row one can choose between 190 THB or the lottery. The switching row from the lottery to the safe payoff designates individuals' risk attitude. A risk-neutral subject would switch when the EV of the lottery equals to the safe payoff; that is row 16. As a result, rows below 16 indicate risk-averse behavior while risk-loving individuals would favor rows above 16.

Alternatively to looking at the switching row, one could study the effects of experimental conditions in terms of the constant relative risk aversion (CRRA) characterization, employing an interval regression model. The CRRA utility function is defined as:

$$U(y) = (y^{1-r})/(1-r)$$

where r is the CRRA coefficient. The dependent variable in the interval regression model is the CRRA interval that subjects implicitly choose when they switch from lottery to safe payoff. For each row in Table 1, we can calculate the implied bounds on the CRRA coefficient. If we assume a constant risk aversion (CRRA) utility function the first row becomes:

$$0.5(\frac{300^{(1-r)}}{1-r}) - (\frac{10^{(1-r)}}{1-r}) \geq 0$$

Where r is the Arrow -Pratt coefficient of relative risk aversion defined as:

$$-\frac{YU''(Y)}{U'(Y)}$$

The solution is 0.796. Therefore we conclude that the solution to the above inequality is $r = 0.796$. In the second row, again, if we assume a constant risk aversion (CRRA) utility function this becomes:

$$0.5(\frac{300^{(1-r)}}{1-r}) - (\frac{20^{(1-r)}}{1-r}) \geq 0$$

The solution is 0.744. We find that an estimate for the coefficient of relative risk aversion for a person who chose Option B is such that $0.796 \leq r \leq 0.744$. We use the same method to estimate the coefficient of relative risk aversion for all choices. Thus, for example, a subject that made 5 lottery choices and then switched to the safe payoff alternatives would have revealed a CRRA interval between 0.613 and 0.656.

With regard to the non-incentivized risk task, the risk survey has been collected in the household survey repeatedly in 2008, 2010, and 2013. The risk survey item is very popular and has been widely used in the field. The general risk question is extracted from the German socio-economic panel study (SOEP) and asks "Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?" Responses are given on an 11-point Likert scale where the value 0 means "unwilling to take risks" and the value 10 means "fully prepared to take risks".

The commonly known drawback from the survey items, however, is that it remains doubtful whether self-reported personal attitudes are behaviorally meaningful because it does not offer an incentive-compatible measure of risk attitudes thus facilitating self-serving biases and inattention that distorts risk attitude (Camerer and Hogarth, 1999). The second drawback is that it does not allow us to make specific inferences about the level of individuals risk attitude as the distinction between risk-neutrality/seeking is not clear. We focus our analysis of changes in risk attitude, therefore, on the experimental measure. We take, however, the risk survey item to corroborate our experimental results.

1.2.3 Experimental Results in the Cross-Section

The cross-sectional experimental outcome for each wave is illustrated in Figure 1. For the experimental result, it is evident that the majority of respondents chose rows below 16, thus explaining the left-skewed shape of the histograms. Second, there is a larger peak at row eleven, i.e. where people seem to be attracted by the safe payoff of a round figure which is 100 Baht. Another smaller peak occurs at row 16, i.e. where the safe payoff exactly matches the EV of the lottery. Third, there is another large peak at the very right, i.e. in the domain of risk-loving answers where subjects prefer a lottery with outcomes 0 or 300 Baht to safe payoffs of 190 Baht. Comparing all three together, there seems to be an increase in the variance of responses over the waves, particularly between wave 1 and 2 signifying the largest increase in risk-loving responses.

1.2.4 Experimental Results over Time

With respect to individual changes over time, Figure 2 displays the data for the within sample comparisons that our experimental design is constructed to allow. The values show the difference between the mid-point of the elicited CRRA interval from the repeated respondents experiment at different points in time. If there were no changes in elicited CRRA, then the data point underlying the histogram in Figure 2 would be zero. If the CRRA had increased in the later series, the data point would have been positive. The differences in relative risk aversion for all three waves illustrated in Figure 2 do have a tendency to be negative, indicating a within-sample difference in risk aversion towards risk neutrality and risk loving attitude -– confirming Figure 1. [1]

Pooling over three years, we find a between-subject standard deviation of 0.37 which is smaller than the within-subject standard deviation of 0.46. This shows that the variation in response due to time is greater than the variation due to individual heterogeneity in subjective preferences.

We can also look at the correlation between CRRA values elicited at different points in time. In the first regression the dependent variable is the CRRA in wave 3 and the independent variable is the CRRA in wave 1. The results show that there is a significant correlation between the elicited CRRA values in 2013 and those values in 2008 with a correlation coefficient of 0.093 (p-value=0.05). This is true for the entire sample. If we only take the respondents that could be re-interviewed (N=418), the correlation coefficient is slightly higher; the coefficient

[1] The distribution of the elicited within-subject difference of the CRRA between 2008 and 2013 interval size is slightly left skewed with a mean of -0.27 and a median of -0.09. The inter-quartile range is between -0.677 and 0.09. The distribution of the elicited within-subject difference of the CRRA between 2010 and 2013 interval size is also slightly left skewed with a mean of -0.09 and a median of 0. The inter-quartile range is between -0.50 and 0.27. The distribution of the elicited within-subject difference of the CRRA between 2008 and 2010 interval size is also slightly left skewed with a mean of -0.16 and a median of -0.093. The inter-quartile range is between -0.43 and 0.135.

of 0.144 is significantly different from zero with a p-value of 0.06.[2]

Evidence with regard to the stability of risk preferences suggests relative low correlation of risk attitudes over time between 0.14 to 0.17 depending on the time interval. Our correlation results are in line with the study of Guiso et al. (2014) and Sahm (2012) where correlations with over 100 observations are 0.18 and 0.13 respectively.

1.3 Possible Determinants of Changes in Risk Aversion

In order to examine possible determinants of individual risk aversion we introduce four groups of variables. Section 1.3.1 displays a large set of socio-demographic and village characteristics and how they are related to risk aversion. Section 1.3.2 demonstrates descriptive of well-being and expectations indicators which are backward and forward looking, respectively. Shocks to the household are presented in Section 1.3.3. Section 1.3.4 discusses possible macroeconomic influences on risk aversion.

1.3.1 Summary Statistics of Individual, Household, and Village Characteristics

In this section we analyze socio-demographic characteristics of our sample population where we distinguish between the repeated sample of 384 respondents who participated in all waves and the total sample of 950 respondents. Table 1.A, Panel A shows summary statistics of the individual and household characteristics of our sample (see Appendix B for variable descriptions). Our sample includes

[2]In the second regression the dependent variable is the CRRA wave 2 and the independent variable is the CRRA wave 1. The coefficient is also here different from 0 and equal to 0.104 (p-value of 0.02). When taking only the repeated observations (N=616), we have coefficient value of 0.173 and a p-value of 0.02. For 2013 and 2010, the coefficient is 0.10 (p-value of 0.03). If only taking respondents from both years (N=516), we have a coefficient that is different from zero and is 0.16 with a p-value of 0.02. If only taking respondents from both years (N=516), we have a coefficient that is different from zero and is 0.16 with a p-value of 0.02.

individuals from rural households who are mostly household heads. On average respondents are 53 years old. Almost 60% of the respondents are female. More than half are engaged into farming activities while few are self-employed. Over the years, we find that the proportion of farmers increases slightly (+9 percentage points), while the number of self-employed individuals decreases (-6 percentage points). This may arise because self-employment is to some degree chosen due to necessity and thus decreases after the 2007/08 crisis. Respondents are typically married, have an average household size of four and low levels of education, i.e. less than six years of schooling. Furthermore, our sample reports to have a good health status which, however, deteriorates slightly over the years. Consumption increases significantly over five years representing a clear improvement in living conditions. Nevertheless, compared to the rest of Thailand, annual average household consumption is still fairly low ($8292.70 PPP). We take logs of consumption in order to come closer to compare relative changes in income than absolute changes. We find substantial differences between full and repeated sample exist along the axes of marital status, age, household size, education, employment, and health status.

Table 1.B, Panel B provides summary statistics of the experimental and non-experimental results over time and for different samples. The average switching row from the lottery to the safe payoff is 7.90 and increases in 2010 to 9.14 which is in line with higher values found in the WTR (Gen). Despite the rise, it still means that the majority of respondents is risk-averse -– corresponding to results typically found in the literature, such as Harrison et al. (2007) for Denmark, Dohmen et al. (2011) for Germany or Hardeweg et al. (2013) for Thailand. By 2013, however, the switching row decreases again, indicating an increase in risk aversion. We find that the repeated sample has a higher average switching row than the full sample.

Table 1.B, Panel C displays descriptive statistics on certain village characteristics. Most importantly, we do not find any significant result on the number of village shocks between repeated and non-repeated sample. The average number

of shocks in 98 villages was 1.48 ranging from 1 to 3 shocks in total. The average number of seasonal workers is 2.07 and does not vary much over time. The number of problems reported by the village head is on average 5.50. The three main problems reported are drug abuse, trafficking, and politics. The number of social and cultural activities among villages is also evenly spread where villages seem to have 6.77 activities on average (i.e. community meeting, festivals following Buddhist traditions etc.). Furthermore, we find that the average distance to the next district capital and to the provincial capital Ubon Ratchathani is 15 km and 62 km, respectively.

1.3.2 Summary Statistics of Well-Being and Optimism Indicators

Table 2, Panel A displays the subjective assessment of the respondent's well-being. The fact that a certain amount of time passed means that there may have been changes in the subjective well-being that could have an effect on elicited risk attitudes. In each wave we asked subjects to respond four questions about the development of the state of their household finance in the short term and long term: "Do you think your household is better off than a year /5 years ago?" and "Do you think your household will be better off in the next year / in 5 years?"

The interpretation of self-assessed well-being indicators have three dimensions: first, it is clearly a subjective measure being influenced by the personality, such as the optimism of the person. Second, individual circumstances and expectations may play a role, such that a finished education promises better income in the future etc. Third, the overall environment will influence the assessment, for example, an improving macroeconomic situation tentatively improves the expectations (Andersen et al., 2008b). From these three dimensions, the second and third dimension mentioned above are already partially captured by the individual characteristics, the shocks and the wave dummies (controlling for changes in the macroeconomic situation). Therefore, we interpret the responses to these items as indicators for the individual's self-assessment of their current personal situation compared to the past and their expectations about their financial condition in the future.

Higher values indicate a worsening situation. The well-being variable indicates the pessimsic view of the fianncial condition of the household in five years time while the optimism indictates how well the household feels at the time of the survey compared to the past. The decreasing values over the years show that respondents feel that their situation is clearly improving over time whatever specific question or reference point we use. We, again, find differences between the repeated and full sample.

1.3.3 Summary Statistics of Microeconomic Shocks

The third group of variables, documented in Table 2, Panel B informs about the number, kind, and magnitude of shocks that households were exposed to before the respective survey. We distinguish between four kinds of shocks: demographic (e.g. illness, death), social (e.g. theft, law suit, conflict with neighbors), agricultural (e.g. drought, flood) and economic shock (e.g. increase in price of inputs, collapse of business). We also distinguish among these types of shocks between the number of covariate and idiosyncratic shocks. Idiosyncratic shocks are assumed to be uncorrelated across households within a community and should therefore be insurable by, for example, informal mutal insurance mechanisms within communities. Covariate shocks on the other hand are correlated across households within the same community and informal insurance mechanism within communities should therefore breakdown during covariate shocks. To differentiate between idiosyncratic and covariate shocks we ask the respondent to estimate the impact of the particular household shock on others. Response categories are no other household, some other households, or most other households in village, district, province, or country. We code shocks of the first two categories as idiosyncratic shocks and the last four as covariate.

In accordance with the income and well-being measures, we see that the number of adverse shocks is decreasing over time. This applies in particular to agricultural and economic shocks, and is thus also reflected in the decreasing number of covariate shocks, whereas idiosyncratic shocks do not seem to decrease in number to that extent.

Another important aspect is to distinguish between covariate shocks and general changes in the macroeconomic environment. Covariate shocks are significantly different from the macroeconomic environment since the spread of distribution of covariate shocks on households is highly unequal. A large share of respondents (60%) did not experience any covariate shocks, while 5% experienced up to 6 covariate shocks over the three waves. Hence, while covariate shocks can be regarded as village-level shocks assigned to each household within a community, changes in the macroeconomic environment should affect all households within a district or a province and, in the worst case, even across regions and other countries. Few studies have attempted to estimate the relative importance of covariate shocks (see e.g Dercon and Krishnan, 2000). Their estimation show that covariate shocks have a larger and more significant impact on household's consumption and vulnerabiliy than idiosyncratic shocks.

1.3.4 Macroeconomic Environment and Risk Aversion

The macroeconomic environment of Thailand between 2007 and 2013 is characterized by unusual high volatility. Compared to other emerging economies with largely functioning institutions, the long run growth rate of Thailand has been high compared to advanced economies. Due to nowadays low population growth (below 1% p.a.); most of this high growth is translated into high per capita income growth.

Turing to concrete numbers, the 4.2% GDP growth rate in the year 2007 is a good starting point, although it is somewhat below the long-term average. Different from many other emerging economies, Thailand could not shield its economy against the recent world-wide financial and economic crisis which is reflected in the respective recession growth rates of 1.8% (2008) and even -3.0% (2009). The bottom of the crisis occurs around the second and third quarter in 2009. Thereafter, a new strong upswing takes place leading to GDP growth of 7.8% in 2010. The following year is impacted by heavy flooding in the central regions of Thailand where industrial production is located and thus growth comes down to 0.01% in 2011. Again, a strong recovery follows with 6.5% growth in 2012, and some normalization with 3.1% in 2013.

The rural areas in North-east Thailand, where agricultural production dominates, are of course affected by these developments. Besides spillovers via demand for agricultural products, financial transfers within families and even within-country migration, there are also heavy price changes for agricultural products. These occur largely in parallel to the overall macroeconomic situation, i.e. prices crash in 2008 and explode in 2010.

Overall, our first wave in spring 2008 is overshadowed by a massively declining macroeconomic environment, whereas the wave in 2010 is influenced by the strong recovery and good expectations. This macro pattern seems to be reflected in the median responses to the risk aversion experiment when looking at the results in Figure 1 which shows an exceptional change in 2010. The negative environment and outlook in 2008 may lead to more risk aversion, the optimism in 2010 to less risk aversion, while the improved situation in 2013 a return to risk aversion levels compared to 2008.

1.4 Test for Attrition and Stability of Risk Preferences

Before coming to our main analysis, Section 1.4.1 documents tests of attrition. Section 1.4.2 presents results on the impact of attrition bias on risk preferences.

1.4.1 Test for Attrition

Of the study with 950 respondents in 2008 we are able to measure risk preferences between 2008 and 2013 with 384 individuals only. A panel for these 384 individuals is created, allowing us to correlate individual preferences over time and to observe whether changes in income, unemployment status, education, or family composition correlate with changes in preferences.

The panel of returnees is clearly a selected sample. For the purposes of this study, selective attrition would be problematic if it was correlated with the stability of risk preferences. Since we could not control who would accept our invitation to participate again, there is the potential for sample selection through attrition. More formally, attrition bias will occur if the error term in the equation of interest is correlated with the error term in the selection or attrition equation. In this respect, attrition bias is model specific as the correlation between the error terms will depend on the precise specification of the model. Prior to calculating a selection model, which relies on identifying a set of instrumental variables, z, which are correlated with attrition but not with ε (Heckman, 1979), it is first essential to test whether attrition in a panel data model is random. We implement the approach by Fitzgerald et al. (1998) and Becketti et al. (1988).

The simplest test for whether attrition is random is to estimate a probit model in which the dependent variable takes the value one for households which drop out of the sample after the first wave. We test for non-randomness by regressing the repetition of the experiment with socio-demographic (i.e. age, employment status, household size), some village level (i.e. number of shocks, number of village inhabitants, distance to town) and specific experimental characteristics (i.e. payoffs) in the previous waves[3]. These variables with the exception of distance to town are included in a selection type model, as they are correlated with both attrition and

[3] We also test whether the last switching row is correlated with the winnings in the previous waves and found no statistically significant results.

stability of preferences. The pseudo R-squared from the attrition probit suggests that baseline variables explain about 5% of panel attrition between 2008 and 2013 (available upon request). Thus, it still leaves some 95% of attrition as unexplained. The z-statistics and p-values show that 6 out of the 12 variables in the attrition probit are statistically different from zero.

The variable that is a high significant predictor of attrition is the age of the respondent (p-value=0.01). The younger the respondents, the more likely are they to attrit which might be related to the higher possibility of moving easier between jobs or migrating to the city. Further variables include the household size — an increase in household size makes it more likely for the respondent to attrit at the 1% significance level. Furthermore more income, worsening of health, and distance to town are explaining attrition. Being a farmer reduces the probability of attriting. The resulting Chi-squared statistic of 39.26 with 12 degrees of freedom indicates these variables are jointly statistically different from zero at the highest level of significance (the p-value=0.000), so we can conclude these are significant predictors of attrition. Given that both the standard tests indicate that attrition for the expenditure model is non-random, we can proceed using a Heckman selection model.

For this, we need to find a variable being orthogonal to stability of preferences but correlated to the probability of attrition. It is helpful that we have the information of distance to the nearest district town that might be orthogonal to risk preferences but have an impact on attrition due to migration. Thus, we take the distance from the respective interviewed village to the nearest district town (average distance in 15 km). Table 1.B, Panel C indicates that distance does indeed correlate with attrition. Individuals, who have a higher distance to the nearest town, are more likely to repeat the experiment and survey. Thus, respondents who live in remote areas are more likely to repeat the experiment than those living close to nearest city – which might be driven due to migration. We do not find the same result for the distance to Ubon which on average is 62 km. Both full sample and repeated sample are equally likely to repeat given the distance to Ubon. Hence, we can use distance as an exogenous determinant of sample attrition and check for biased estimations of risk preferences due to sample attrition in the next analysis.

1.4.2 Stability of Preferences with Socio-Economic Characteristics and Attrition

In Section 1.2.4 we saw that the correlation between 2008 and 2013 of the experimental results was 0.14 and highly significant. However, such stability in risk attitude may be caused by selective sample attrition. That is, it may be the case that only individuals with stable preferences elect to participate again in the experimental study. It may also be the case that socio-demographic characteristics reduce the five-year correlation in preferences.

In order to establish a baseline for evaluating the influence of demographics and selection on the stability of risk attitude, column (1) of Table 3 estimates pooled-OLS regressions of CRRA in 2013 on CRRA in 2008 with a constant. The procedure is similar to Andersen et al. (2008b). Column (2) includes socio-economic characteristics included in Table 1. Preferences remain significantly correlated over time. Column (3) shows estimations of the Heckman 2-step procedure to control for selective attrition from the study. We do not consider a permanent unit non-response since some individuals who have been missing in wave 2 reappeared for the last wave. In other words, we treat the panel as a cross-section and employ the maximum-likelihood estimator. This means that our sample is increased from 384 who repeat our experiment in all three waves to 471 respondents who participated in 2008 and were re-interviewed in 2013.

As a first step, the inverse mills ratio is generated from the probit regression of non-random attrition including the distance to the next district town, number of inhabitants, and the number of village shocks. Under the assumption that distance is orthogonal to preference stability, column (3) identifies the temporal correlation in preferences controlling for stability-driven attrition. We regress CRRA2013 on CRRA2008 including the newly-generated inverse mills ratio as a sampling weight. We see that controlling for sample attrition does little to the estimated correlation.

Looking at the three exogenous characteristics in column (2) and (3) – gender, age, and height (see Dohmen et al., 2011) – we find that using the Heckman-selection model yields different results. Gender seems to be significant and positively correlated with risk aversion which is in line with the literature documenting differences in the risk preferences of men and women (Croson and Gneezy, 2009). Age seems to be significant in column (2) but the coefficient sign is theoretically surprising and we will indeed see later, that this coefficient is not robust – it basically captures the effect from the later waves, which are not controlled for here. Nevertheless, the results so far suggest relative low correlation of risk preferences over time even after taking into account individual characteristics and sample attrition.

1.5 Identification Strategy

In this section, we present our identification strategy in order to demonstrate the impact of four groups of variables on changes in risk preferences: these groups are individual characteristics, individual well-being indicators, the experience of shocks, and the macroeconomic environment. Section 1.5.1 displays the econometric model to investigate systematic determinants of changes of risk preferences. Section 1.5.2 discusses the possible impact of selection bias and presents results.

1.5.1 Econometric Model

The novelty of our study is that we use panel data. Most research on time-varying risk aversion is based on cross-sectional data (e.g. Binswanger, 1980; Donkers et al., 2001; Harrison et al., 2010) and hence is subject to the usual limitations associated with such data.

By presenting evidence from a representative panel data set, we further can infer the direction of causation - do shocks lead to higher risk aversion or higher risk aversion to higher shocks? Albeit it is difficult to solve the endogeneity problem fully, we argue that shocks rather happen unexpected for all individuals alike. Our basic idea is similar to difference-in-difference: we capture the effect of microeconomic shocks and worsening of well-being by comparing individuals with zero exposure and non-zero exposure assuming that the level of risk aversion would have been the same in the absence of the microeconomic shocks. Further, with repeated observations for the same individual, it becomes possible to control for unobserved time-invariant individual specific effects.

Our identification strategy exploits the determinants of changes in risk preferences, while controlling for four groups of variables - individual characteristics, subjective assessment of well-beings, shocks, and macroeconomic environment using the individual fixed effects model. More formally, we estimate:

$$Y_{i,t} = \alpha t + \beta 1 Shock_{i,t} + \beta 1 Wellbeing_{i,t} + \beta 3 X_{i,t} + \pi W_i + u_{i,t}$$

where αt is the time effect, and $Y_{i,t}$ is the measure of risk preferences using a constant relative risk aversion interval computed from the switching row of individual i at time t. $X_{i,t}$ are observed time-varying individual characteristics. The vectors of controls include age, height, education, marital status, household size, health status and log per capita consumption. πW_i captures any time-invariant unobserved heterogeneity and $u_{i,t}$ is the error term. The coefficient αt captures the time-varying changes on risk attitudes, $\beta 1$ measures the effect of micro-economic shocks on risk preferences while $\beta 1$ measures the effect of changes in well-being on the shift in distribution of risk preferences. We expect a positive coefficient of microeconomic shocks and well-being on the CRRA interval, i.e. the more shocks an individual is exposed over the time, the higher the risk aversion.

We employ the fixed effects estimation in equation (1) to remove time and individual fixed effects. The Hausman test statistic of 35.25 leads to the rejection of the model without fixed effects. Hence the fixed effects estimator is superior compared to the random effects model. In the main part of our analysis we will use fixed effects estimation to identify possible determinants of changes in risk aversion. The reduced-form of the model looks as follows:

$$\begin{aligned} Y_{i,t} - \bar{Y}_i = \alpha t - \overline{\alpha}_i + \beta 1(Shock_{i,t} - \overline{Shock}_i) + \beta 1(Wellbeing_{i,t} - \overline{Wellbeing}_i) \\ + \beta 3(X_{i,t} - \overline{X}_i) + (u_{i,t} - \overline{ui}) \end{aligned} \tag{1.1}$$

A major econometric issue is the possible presence of unobserved fixed effects πW_i that challenges a causal interpretation of results. It can well be the case that the basic part of risk preference $Y_{i,t}$ is driven by some unobserved individual characteristics such as physical and mental stress tolerance. At the same time, these unobserved individual characteristics could be correlated with $X_{i,t}$ through factors such as residential sorting. In order to be able to infer a clear causal relationship between the exposure of shocks and changes in risk aversion, we control whether selective migration undermines our identifying assumption.

1.5.2 Selection Bias

There are two principal issues confronting attempts to establish the causal effect of shocks on economic preferences using the panel data. The first is attrition bias where individuals' decision to repeat the experiment is related to their risk preferences. In Sections 1.4.1 and 1.4.2 we account for attrition bias. Second, individuals may locate according to their preferences – hence our data may suffer from selection bias. In our case risk-averse individuals may select areas which are less prone to shocks. In other words, individuals who live in villages that experienced a number of self-reported microeconomic shocks in the past 5 years may be different from individuals who live in a village that did not experience such shocks. For example, wealthier individuals may choose to live in villages with less exposure to shocks and hence are more likely to choose the riskier option because risk–seeking attitude and wealth are positively correlated in the literature. Hence the causal effect between shocks and risk may be distorted due to the wealth effect.

In the same manner, one can argue that villages that experienced a great number of microeconomic shocks in the past 5 years might be different from villages that did not experience any number of shocks. For instance, villages that have more shocks may have less provision of public goods; infrastructure etc., again causing a negative correlation between number of shocks and risk aversion that is not causal. We follow the approach by Cameron and Shah (2015) in order to control for selection bias.

Table 4 shows – pooling across villages – that there is no significant difference between villages that experienced a shocks and one that did not. In Panel A, we provide village level characteristics and see whether villages that experienced shocks are significantly different from villages that did not have a shock in the past 5 years. Although, a considerable time has passed by allowing changes to happen in a village after a shock, we do not find any difference for variables that are plausibly exogenous to exposure to shocks. The number of social and cultural activities and well as the report of the three major problems seem to be balanced across villages over time. Furthermore, we find that non-shock and has-a-shock-village provide the same amount of water to the households. There seem to be also the same amount of nurseries and banks in villages with and without shocks.

Hence, our villages selected seem to be balanced on average.

Panel B presents the mean and standard deviation of individuals who experienced or did not experience a shock in the past 5 years. Most importantly, we find no significant difference in wealth or consumption between individuals who experienced and did not experience a shock in the past 5 years which gives us an indication that the wealth effects on changes in risk attitude may be negligible. We still test for this in later analysis more thoroughly. We do, however, find a differences in the exposure of shocks in terms of gender (women seem to be more vulnerable to shocks), age, household size (interestingly, more shocks result in a greater number of household size) and health (those with a shock experience report higher values of sickness). Table 4, thus, provides evidence in support of our identifying assumption.

To conclude with, we find that attrition – albeit non-random- is unrelated to individual's risk attitude (Table 3). We also do not find that our result might suffer from selection bias; it does not seem that households decide to live in villages where the likelihood of experiencing a shock is lower based on their preferences. Since we can exclude the two possible confounding factors – attrition and selection bias – we can argue for a causal interpretation of the impact of the number of shocks and changes in the levels of risk preferences.

1.6 Empirical Results

In order to examine the possible determinants of individual risk aversion we proceed in three steps. Section 1.6.1 presents the most important determinants of risk aversion. Section 1.6.2 looks into the income effects of changes in risk aversion and Section 1.6.3 presents the determinants of changes in risk aversion for a subsample.

1.6.1 Estimates on the Determinants of Risk Rreferences

Table 5 reports the results of the within-estimation of the regression reported in equation (1). We present the results from simple fixed effects regression model where the dependent variable is the constant relative risk aversion interval. All specifications allow for clustering of standard errors at the village level. In addition to the time effects, we also include district fixed effects to control for any potential differences at the district level which might affect our results such as public goods provisions, government programs and/or geographic differences. Column (1) includes a number of individual characteristics including the number of microeconomic shocks and well-being indicators that the respondents report excluding time and district fixed effects. Note that our within estimator regression automatically nets out the influence of time-invariant characteristics, which in our case is gender and height since we only work with those respondents with whom we have repeated the experiment in all three waves. In column (2), we run the same regression as in column (1) including a time dummy and in column (3) we include both district and time dummies [4].

In column (1) our estimate shows that large changes in levels of risk aversion are not systematically related to changes in levels of socio-demographic characteristics of our respondents with the exception of age. Higher age is related to lower degrees of risk aversion. Interestingly, we find the elasticities of consumption to be small and insignificant, supporting the CRRA assumption which is similar to the findings of Chiappori and Paiella (2011) and Guiso et al. (2014). Instead, we observe that both optimism indicators explain variations in risk attitude. In column (2) we see, however, in terms of socio-economic characteristics that age becomes insignificant while marital status becomes significant which is not very robust to

[4] Correlations between optimism and well-being indicators are 0.09 (p-value=0.01). Correlations between the number of shocks and well-being indicators are 0.07 (p-value=0.01). Correlations between different kinds of shocks is highest between demographic and social shocks are 0.07 (p-value=0.01). It is lower between all other shocks.

the inclusion of district fixed effects in our model. Most importantly, we find that including time dummies, increases the predictive power of the number of shocks and the perceived well-being indicators which seem to be consistently correlated with higher levels of risk aversion. In other words, more pessimistic views about the future and more exposure to shocks in the last five years, significantly increase the average levels of risk aversion.

The coefficient of changes in aggregate levels of risk aversion is higher considering the pessimistic outlook of the future (0.058, p-value=0.05). The year coefficients are all comparisons with 2008 and are all negative. Thus, other things being equal, the risk aversion of our respondents decreases significantly in 2010 compared to 2008. Yet it is statistically insignificant in explaining temporal instability in risk aversion. Finally, given the number of combinations we tried, we still find that the 'classical' determinants of risk aversion almost virtually no explanatory power across all models in conjunction with the general macroeconomic environment captured by the time dummy. Instead, we find number of shocks and well-being indicators to be crucially explaining temporal instability of risk preferences. Our results and its magnitude remain statistically significant even after controlling for district fixed effects (column 3). Given these findings, it is legitimate to ask which kinds of shocks might affect the risk aversion of the respondents.

1.6.2 Impact of Different Kinds of Shocks on Changes in Risk Aversion

In Table 6 we further investigate the effect of shocks on changes in risk aversion. In column (1) we include all kinds of shocks with different magnitudes in our regression to investigate whether and in which way the consideration of different types of shock experiences contributes to changes in individual risk attitudes. We again use the within estimator with district and year fixed effects. The most coefficients of individual characteristics are still insignificant. We find that different kinds of shocks differently affect changes in risk aversion. While agricultural, social, and economic shocks significantly increase risk aversion throughout all estimations, the effect of demographic shocks seems to have a converse effect. The inclusion of "household member joined the household" through marriages for instance could result into a decreased risk aversion because it involves the addition of a member in the household and potentially increases and diversifies the contribution to household finances. The result for demographic shocks, however, is statistically insignificant. The strongest effect for an increase in risk aversion seems to come from economic shocks with a magnitude of 0.13 and p-value of 0.000. The high within R-squared of 0.177 demonstrates high response variability.

Column (2) investigates the impact of the magnitude of shocks on changes in risk attitude because the severity of shocks may also have an impact on risk attitude. We rely on respondents who classify the severity of shocks they were exposed to as having low, medium, and high impact. We find that coefficients on high impact shocks significantly increases risk aversion with a coefficient of 0.083 and p-value of 0.000. It is robust irrespective whether we encompass time and district fixed effects in our model.

In column (3), we see that the signs indicate an increase in risk aversion in the case of idiosyncratic and covariate shocks. Results are, however, insignificant, thus neither idiosyncratic nor covariate shocks seem to explain better variations within individual's risk aversion which may be due to interdependencies between shock categories that could distort the result. Similar to Table 5, we find that time dummies – indicating macroeconomic shocks – are statistically insignificant.

The main result for this study is that the use of panel data and fixed effects models corroborates the hypothesis of large negative effect of various kinds of microeconomic shocks on risk aversion. Overall, we find that variation in risk attitudes is not significantly correlated with any of the exogenous variables for risk aversion such as female, age, and height, and also not with family composition such as household size or income. Instead, we find that individuals' perceived wellbeing and the number of shocks they experience have detrimental effects on changes in risk aversion. To summarize, it seems that the number of shocks, specifically high impact economic shocks in conjunction with a deterioration of well-being, a pessimistic outlook of the future seem to explain changes in the level of risk aversion.

1.6.3 Disentangling Income Effect and Shocks on Changes in Risk Aversion

One possible interpretation of our result is that the changes in risk preferences are driven by changes in income or wealth that accompany unanticipated micro-economic shocks particularly in the presence of imperfections in the credit and insurance market (Jappelli and Pistaferri, 2010). Gertler and Gruber (2002), for instance, look at the effect of income shocks arising from major illness on consumption in Indonesia. They find that whereas people smooth out the effect of minor illnesses well (which could be interpreted as transitory shocks, or anticipated events), they experience considerably more difficulty in smoothing the impact of major illnesses (which could be interpreted as unanticipated event). Given major implications of shocks on income, it is important to estimate whether the negative experience of a shock results in a reduction of their income opportunities and hence a change in their respective risk preference.

In our estimations in Section 1.6.1, we control for income/consumption effects. Table 5 and 6 show that consumption is not statistically significant. Instead, we find that micro-economic shocks are a major determinant in explaining variations in risk aversion. However, to examine the role played by income changes due to shocks more closely, we include further variables from the household survey. In our dataset, households were also asked to report the value of asset lost due to the shocks they experienced as well as the amount of financial aid received (if any).

The reported average asset loss due to micro-economic shocks is on average 3% for all households in the past 5 years.

Table 7 runs the same regression as Table 5, using the number of economic shocks in addition to socio-demographic variables since it was the statistically significant shocks that determined changes in risk aversion. We also include additional controls for the log per capita consumption, namely asset loss due to shocks and the amount of financial assistance received. This allows us to examine if the number of economic shocks is still a major determinant after controlling for both a positive income shocks due to financial assistance and negative income shocks due to a loss in assets. As anticipated, the higher the levels of asset loss, the higher is the level of risk aversion while higher financial assistance is negatively correlated to risk aversion. Both, however, are statistically insignificant.

In contrast, throughout columns (1) to (3), we see that the increasing number of economic shocks occurred in the household is still significantly related to risk-averse behavior. The magnitude of the negative effect of economic shocks on risk aversion remains highly significant at the p-value of .050 while its magnitude decreases slightly by including time and district fixed effects (from 0.14 to 0.13). Interestingly, adding further controls enhances the significant relationship between schooling and risk aversion. Higher levels of schooling are negatively correlated with higher levels of risk aversion. Dohmen et al. (2011) found that having highly educated parents diminishes risk aversion. In addition, it seems that both time dummies are statistically significant in this model. This means that the expected value of the risk aversion is constant during the first period, then decreases significantly (p-value=0.01) between the first and the second period and also between the second and the third period (p-value=0.05). This corresponds to the summary statistics on the average switching row in Table 1.B but overall improvement of the financial situation after 2010. Hence, in this model, it seems that the macroeconomic variations seems to have an impact on changes in risk aversion.

Noticeable is also an increase in the within R-squared that is higher than in Table 5 and 6, indicating that the inclusion of asset loss due to shocks and remittances increases the explanatory power of our model. The bottom line from Table 7, however, remains that households that were severly affected by economic shocks - given they have lost assets due to shocks - still become more risk-averse over time. That is, changes in income and wealth due to shocks do not fully explain the more risk-averse behavior of households that experienced micro-economic shocks.

1.6.4 Impact of Shocks on Changes in Risk Preferences (Subpopulation)

In Table 8 we run the same regression as in Table 6, however, examining the magnitude of shocks on certain subpopulations, i.e. whether microeconomic shocks are more prevalent for those having lower income and lower insurance. In Panel A, we split our sample between higher and lower income groups at the median. We find that an increase in the number of economic shocks for all those who repeated the experiment is associated with significantly greater risk aversion over time for people below median income, controlling for socio-demographic characteristics. Interestingly, the coefficient on the magnitude of economic shock on changes in risk aversion is stronger for higher income groups (0.22, p-value=0.000) compared to lower income groups (0.18, p-value=0.050). An explanation is that higher income groups also report higher income and asset loss due to shocks compared to lower income groups. Hence income levels might have an impact on the magnitude of the economic shock.

In Panel B, we make another split and look at the sample who report having fewer number of insurances than the median and those having lower number of insurances. We find that respondents' risk aversion changes more significantly for those who report having lower numbers of insurances than the median. We find that having lower amounts of insurances makes them more vulnerable not only to economic but also to agricultural shocks and it is statistically significant at the 10% significance level. The magnitude of economic shocks on changes in risk attitude for those having fewer insurances is also higher compared to those with higher number of insurances. The magnitude of the number of shocks one experience on levels of risk aversion is even higher compared to Table 6. Looking at the diminishing magnitude of agricultural shocks, we find that already having more than two insurances renders agricultural shocks insignificant in explaining changes in risk aversion.

This result provides further insight into the effectiveness of insurances in reducing economic insecurity and the impact of negative income shocks in developing countries. Insurance instruments can be one of many options in managing risks of unanticipated covariate and idiosyncratic shocks, thus reducing vulnerability. Overall, we find that having low income and fewer insurances increase the influence of microeconomic shocks affecting changes of risk preferences.

1.7 Robustness

In this section, we provide further sensitivity analyses. Instead of using the experimental result, we scrutinize determinants of changes in risk aversion using the survey item in Section 1.7.2. In Section 1.7.3, we alter the dependent variable by using directly the switching row of the experimental results, instead of the constant relative risk aversion interval. In Section 1.7.4 we test whether results are driven by outliers.

1.7.1 Changes of Risk Attitudes Over Time

While we focus in the Section 1.2.2 on the within-subject differences using the experiment, we can extend our analysis of temporal stability of preferences using a psychometric approach concerning the inter-class correlation of responses of all the risk preference items over time. The analysis of reliability has also been used by Beauchamp et al. (2013). The reliability analysis can be thought as the proportion of the total variance that is explained by subjects, analogously to the coefficient of determination R-squared in linear regression models. In other words, it refers to the consistency of individuals' responses to an instrument across occasions under the assumption of measurement error.

An important concept of in psychometric reliability analysis is the test-retest reliability (Rabe-Hesketh and Skrondal, 2012). We employ the test-retest reliability of risk attitudes using maximum-likelihood estimation. For the experimental measure, we find that the overall population mean β is 0.2832. The estimate of the between subject standard deviation is $\sqrt{\theta}$ is 0.1795 and the within standard deviation is $\sqrt{\theta}$ is 0.567. The interclass-correlation is given by rho, which is 0.09. Because of the low interclass-correlation, we conclude an increase in variance over the waves. Hence, answers over different times demonstrate not much reliability or stability at least in terms of the experiment (Results are available upon request).

Using standard regression results and the interclass-correlation methods to measure stability of risk preferences, we find unequivocal evidence for large degree of variations in individuals risk attitudes over five years. Between the first and the second wave, we have a signficant decrease in risk aversion. By the last wave, it seems that responses have fallen back to similar values as in the first wave. This confirms results in Figure 1 and Figure 2.

1.7.2 Determinants of Changes in Risk Attitude using Survey Item

In this section we investigate changes in individual risk attitudes, measured by the 11-point Likert risk item. One argument that has been put forward by Dohmen et al. (2011) is that general willingness to take risk item provides a behaviorally meaningful alternative to experimental measures and can be generally applied in many dimensions of risky behavior. The survey-based measure of risk attitude was validated by an incentive compatible experiment using data from 900 participants in rural Thailand in 2008 (Hardeweg et al., 2013). We replicate the results of Hardeweg et al. (2013) and find that the survey measure of risk attitude predicts risky behavior in the same way as the risk experiment (see Appendix C).

After confirming the behavioral validity of the survey instrument, we proceed to provide some descriptive statistics about the cross-sectional results of the survey item afterwards estimate the same specification used in Section 1.5.1. Figure 3 displays cross-sectional results for the non-incentivized survey item. Comparing Figure 1 with Figure 3, we find a similar, albeit slightly different picture. Figure 3 shows the histogram of the willingness to take risk survey item. We find that the bulk of the sample choose the middle category of 5; a result which is similar to the study of Dohmen et al. (2011). This is the case for the first and the second wave. However, in 2013, we find a shift in the distribution with a spike at the right hand side. Not surprisingly, the mean willingness to take risk is shifted from 4.53 to 6.80 (see Table 1). Full and repeated sample also have almost identical average levels of risk aversion for the WTR (Gen). In the next step, we will investigate changes in risk aversion using a fixed effects model and the same covariates as for the experiment (see Table 5).

Similar to Table 5, we find in Table 9 that increasing number of shocks decrease the willingness to take risk over time. This also holds if respondents fell less well-off compared to the past five years which is statistically significant at the 1% significance level even after taking into account time and district fixed effects. The magnitude of the number of shocks is even higher compared to Table 5. We, however, do not find any impact of pessimism on the willingness to take risk. Another deviation from Table 5 is that the willingness to take risk is positively associated with the age of the respondent. Nevertheless, we find that the number of shocks, as well as the well-being measure significantly explains variations in the willingness to take risk item over time in the same direction using the certainty equivalent task.

Table 10 confirms results in Table 6. Employing the same specification and merely altering the dependent variable yields very similar results. Most importantly, we find that the number of economic shocks is still statistically significant in explaining lower levels of willingness to take risk. In a similar manner, we still find that high impact shocks is negatively associated with higher levels of willingness to take risk (0.25, p-value=0.05). The number of covariate, idiosyncratic shocks and time dummies, however, still remain insignificant.

Evidence provided in this section proves that that the detrimental effect of shocks on risk aversion persists after introducing another risk attitude measure.

1.7.3 Determinants of Changes in Risk Attitude using Switching Row

In the previous analyses, we assume that responses to the experiment follow a CRRA utility function. In the context of utility, risk aversion is equivalent to concavity of the instantaneous utility function, and if one is willing to make particular assumptions about the functional form of utility, it is possible to calculate risk preference in terms of a parameter describing curvature. This calculation uses the assumption that utility is defined only over outcomes in the experiment, rather than over final wealth levels (for a similar approach see Holt and Laury, 2002). Thus using the CRRA as a dependent variable makes strong assumptions.

What we do now is, instead of taking the midpoint of the CRRA interval as the dependent variable, we use the row at which the respondent decided to switch from the lottery to the expected payoff. Hence, we allow risk attitudes to enter non-parametrically. Low values indicate risk aversion which differs from the CRRA where negative values indicates risk-loving behavior.

What can be inferred from Table 11 is that all variables from Table 5 are still robust. Most importantly, experiencing higher levels of shocks over time increase levels of risk aversion. Furthermore, it seems that if respondent's become more pessimistic over the years, they also are more likely to be become more risk averse, i.e. have lower switching rows. However, in contrast to Table 6, changes in marital status seem decrease switching rows. One explanation could be that entering a marriage relationship often corresponds with higher responsibility within the household. This is the only socio-demographic variable that seems to positively to affect risk aversion, however, only in the within estimator without district fixed effects (column (1)-(2)). We also observe here the partial impact of general macroeconomic volatility. It seems that including year dummy 2010 explains now a significant decrease in the levels of risk aversion at the 5% significance level. Thus, it confirms our descriptive statistics showing that in 2010 risk aversion decreased significantly.

In Table 12 we further investigate the effect of different kinds of shocks on changes in risk aversion using the switching rows. With regard to kinds of shocks, Table 12 shows that economic shocks are still significant explaining increases of risk aversion over time. We still find that coefficients on high impact shocks significantly increases risk aversion with a coefficient of -1.32 and p-value of 0.000. Interestingly, we find now that agricultural shocks, covariate shocks as well as macroeconomic volatility – captured by the time dummy – seem to be significant using switching rows instead of the constant relative risk aversion interval.

As noted before, using the CRRA interval relies on strong assumptions. As a way of circumventing the need to make strong assumptions about the form of utility, we use the same regressions of changes in risk aversion as in Table 5 and 6, however, using directly the switching rows from the experiment. This approach implicitly assumes that the experiment measure captures the curvature of the unknown, not predetermined, utility function. In this case we find a strong and significant impact of the time dummy and covariate shocks to explain changes in risk aversion which we are unable to observe using the CRRA as the dependent variable. Using the switching row, we find slightly stronger results of the impact of macroeconomic fluctuations, well-being measure, and microeconomic shocks on changes in risk aversion.

Taken together, the picture emerging from the results in this section suggests there is a relationship between risk aversion and micro-and macroeconomic shocks, although we cannot completely rule out that curvature plays a role in explaining the magnitude of these results on changes in risk aversion.

1.7.4 Determinants of Changes in Risk Attitude using the Upper Bound of CRRA

Using the switching row of the certainty equivalent task allows us to generate household-specific risk-aversion intervals mean values of the upper and lower bound of the CRRA for each switching row. We described the method in more detail in the Section 1.2.2. In this section, we take the upper bound of the risk aversion parameter as the dependent variable to investigate whether our main results in Table 6 could be driven by extreme outliers. We use the upper bound of the interval. We find that our main empirical results are not sensitive to the choice of lower switching rows.

Table 13 replicates results from Table 6 and finds that microeconomic shocks – particularly economic shocks - are still highly significant in explaining variations in risk aversion. Its magnitude, however, is slightly less than in Table 6 which is plausible given the exclusion of the lowest switching row (0.098). We also find that high impact shocks increase levels of risk aversion over time at the highest level of significance. We still not find any relationship between changes in the socio-economic condition of the household that could explain changes in risk aversion. In contrast well-being and optimism indicators are still statistically significant. The dummy of 2010 is weakly negative significant indicating that there is an impact of macroeconomic environment on variations of risk aversion, yet it is not as dominant as microeconomic shocks. This is an interesting result because traditional empirical studies (Townsend, 1994; Jalan and Ravaillon, 1999; Dercon and Krishnan, 2000) show that macroeconomic shocks have a larger and more significant impact on households' livelihood than idiosyncratic or household-level shocks due to informal within-family insurance mechanisms that can alleviate some adverse effects of the latter. Our study, thus, provides a new insight. We do find a marginal effect of macroeconomic environment on changes in risk aversion, yet the effect and magnitude is especially more significant and robust looking at microeconomic household level-shocks.

Overall, we do not find that our main results are driven by extreme outliers. Disentangling the micro-macro effect on changes in risk aversion, we find that microeconomic shocks and well-being indicators are still the major determinants in explaining variations in risk aversion over time.

1.8 Conclusion

Risk and uncertainty play a significant role in almost every important economic decision making. Since people differ in the way they take decisions involving risk and uncertainty and since these differences are often described as differences in risk attitude, understanding individual risk preferences is a prerequisite to understand economic behavior.

There are studies exhibiting influences of socio-demographic characteristics on risk aversion in the cross-section, however, many of them have small samples, quite limited socio-demographic information or do not use incentivized experiments. Our knowledge about risk aversion over time is more sparse, since there are only a few instances of experiments conducted with the same group of people over long time periods.

This research is the first to repeat a standard incentivized experiment eliciting risk aversion with a larger group of people over a period of several years. It attempts to disentangle micro-and macro influences on changes in risk aversion over time. Using panel data, correcting for attrition and selection bias, we try to establish a causal relationship.

In our case two risk measures are conducted with about 900 subjects over a period of five years. Due to attrition of households and unavailability of individuals, we use a sample of 384 respondents who have participated in all three experiments.

First, we describe and analyze the distribution of risk aversion across individuals and changes over time. We find that risk aversion decreases significantly over time. Looking into the determinants of changes in risk aversion, we find that the increasing number of adverse shocks lead to more risk aversion. This applies in particular to shocks, which are regarded as having high impact, being idiosyncratic in nature and occuring in the fields of agriculture and economics. This is particularly pronounced for low income and low insured households.

Second, subjective well-being impacts risk aversion. Risk aversion is positively affected if people either regarding their situation worse compared to the past or expecting a worsening situation in the future. Both cases positively affect risk aversion. These measures may partially include the experience of negative shocks, which is of more objective character, but they also include the individual ability to get along with circumstances and of course the degree of optimism when looking into the future.

Third, the decreasing risk aversion over the five years period indicates, controlled for other influences, that the macroeconomic environment may also play a marginal role. The coefficients on the wave dummies are significant, yet its impact weakens in relation to micro-economic shocks. However, the level changes of the time dummies seems to be in line with the macroeconomic changes.

Overall, this study shows that the risk aversion of 384 people is systematically affected by major changes in living conditions. We identify adverse shocks which increase risk aversion, see how realized or expected worsening in well-being increase risk aversion and find that an improving macroeconomic environment reduces risk aversion. All evidence calls for caution when drawing far reaching conclusions from eliciting individual risk aversion at a single point in time.

Another important aspect that we cannot measure is whether changes in risk attitude are indeed temporal or permanent. It would be interesting to investigate the longevity of the impact of certain shocks on risk aversion further. We have a modest indication that an increase in risk aversion is temporary. Looking at the distribution of risk aversion over years, we find that after the peak in 2010 levels of risk aversion increase again. In psychology literature, Burns et al. (2011) conducted a panel study to understand the trajectory of risk perception amidst the ongoing economic crisis. Their studies showed that while people's risk aversion increased during the crisis, over time, it returned back to baseline level because people adapt to ongoing stimuli. As a result, they hypothesize that risk aversion follows an inverted U-curve – similar to our finding. However, this is only a first insight and a deeper understanding would certainly be desirable to make further inferences on the evolution of risk aversion over a longer time period.

Finally, while we know that risk does not seem to stable over time and it is determined by the increasing number of certain micro-economic shocks, it is interesting to investigate whether changes in risk attitude is also translated into risky behavior (i.e. more risk aversion resulting in the possible reduction of investment-related activities etc.). Future research should carry out deeper investigations about the impact of changes of individual risk attitude on risky behavior.

Figure 1: This figure reports the distribution of the experimental measure of risk aversion in 2008, 2010, and 2013 using a certainty equivalent task. In a nutshell, respondents make 20 decisions between a safe payoff and a lottery, where the lottery remains unchanged but the safe payoff increases steadily from the decision to decision. Further details on the experimental procedure are displayed in Appendix A.

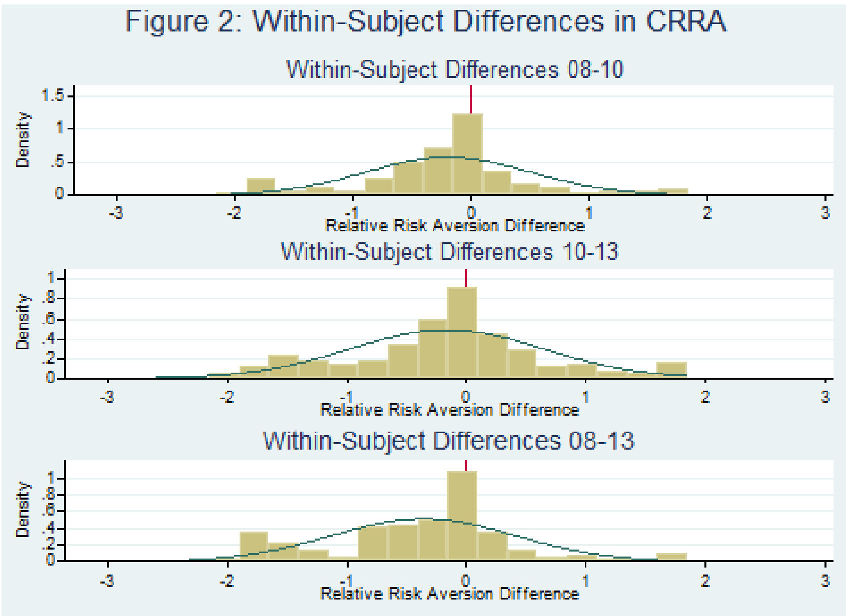


Figure 2: This figure shows the within sample variation in the experimental task. The values show the difference in CRRA values of the switching row between 2008-2013, 2010-2013, and 2008-2010.

Table 1.A: Summary Statistics for Full and Repeated Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Participation in study			Sample in Different Years			
	(Full Sample)	(Repeated Sample)	(T-Test)	(2008)	(2010)	(2013)	(T-Test)
Panel A: Socio-demographics							
Gender	0.58	0.59	0.57	0.57	0.59	0.59	0.23
	(0.49)	(0.49)		(0.50)	(0.49)	(0.49)	
Age	52.36	54.33	0.00	50.75	52.00	54.43	0.00
	(13.35)	(11.84)		(13.02)	(12.84)	(13.97)	
Height (cm)	157.12	157.87	0.18	158.11	158.80	158.20	0.83
	(9.01)	(8.99)		(10.66)	(7.90)	(8.13)	
Farmer	0.63	0.67	0.00	0.60	0.62	0.69	0.00
	(0.48)	(0.47)		(0.49)	(0.49)	(0.46)	
Public Servant	0.03	0.02	0.00	0.03	0.02	0.02	0.16
	(0.16)	(0.12)		(0.18)	(0.15)	(0.15)	
Self-employed	0.09	0.07	0.00	0.11	0.10	0.05	0.00
	(0.29)	(0.26)		(0.32)	(0.30)	(0.22)	
Martial Status	0.83	0.80	0.01	0.82	0.84	0.82	0.75
	(0.38)	(0.40)		(0.38)	(0.37)	(0.39)	
Years of Schooling	5.63	5.27	0.00	5.57	5.63	5.68	0.48
	(3.04)	(2.69)		(3.01)	(2.97)	(3.14)	
Household Size	4.09	3.90	0.00	4.13	4.09	4.04	0.31
	(1.75)	(1.74)		(1.79)	(1.76)	1.72	
Health Status	1.56	1.59	0.08	1.59	1.48	1.61	0.47
	(0.71)	(0.70)		(0.75)	(0.65)	0.73	
Log per capita consumption	7.51	7.50	0.78	7.26	7.69	7.57	0.00
	(0.63)	(0.62)		(0.64)	(0.56)	(0.62)	
Observations	2651	1152		942	900	849	

Notes:This table shows means of individual, household, and village characteristics. Standard deviations are reported in parenthesis. Column (3) shows p-values for T-Tests between Column (1) and Column (2). Column (7) shows p-values for T-tests between columns (6) and (4).

Table 1.B: Summary Statistics for Full and Repeated Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Participation in study			Sample in Different Years			
	(Full Sample)	(Repeated Sample)	(T-Test)	(2008)	(2010)	(2013)	(T-Test)
Panel B: Risk Atttitudes							
Certainty equivalent	7.90	8.36	0.00	6.79	9.14	7.83	0.00
	(6.37)	(6.58)		(5.54)	(5.99)	(7.33)	
Willingess to take risk (WTR)	5.36	5.30	0.38	4.53	4.97	6.80	0.00
	(3.00)	(3.04)		(2.95)	(2.54)	(3.60)	
Panel C: Village Characteristics							
No of village shocks	1.48	1.46	0.64	1.47	1.49	1.47	1.00
	(0.74)	(0.73)		(0.74)	(0.75)	(0.74)	
No of seasonal workers	2.07	1.91	0.21	2.07	2.06	2.07	1.00
	(5.07)	(4.83)		(5.10)	(5.11)	(5.10)	
Major problems	5.50	5.46	0.88	5.49	5.52	5.48	0.96
	(12.74)	(12.58)		(12.71)	(12.81)	(12.71)	
No of social activities	6.77	7.95	0.18	6.76	6.76	6.78	0.96
	(6.33)	(6.37)		(6.34)	(6.33)	(6.34)	
Public Water Provision	0.93	0.93	0.22	0.93	0.93	0.93	0.93
	(0.26)	(0.25)		(0.26)	(0.26)	(0.26)	
Waste Water	1.16	1.14	0.70	1.16	1.15	1.16	1.00
	(0.65)	(0.62)		(0.65)	(0.66)	(0.66)	
Solid waste management	1.38	1.38	0.70	1.39	1.40	1.39	1.00
	(0.55)	(0.54)		(0.55)	(0.55)	(0.55)	
Has a nursery	0.50	0.48	0.06	0.50	0.51	0.51	0.98
	(0.50)	(0.50)		(0.50)	(0.50)	(0.50)	
Has a bank	0.08	0.08	0.74	0.08	0.08	0.09	0.99
	(0.27)	(0.27)		(0.27)	(0.27)	(0.27)	
Distance to district town	14.85	15.65	0.04				
	(9.70)	(9.89)					
Distance to provincial capital	61.71	62.02	0.83				
	(35.97)	(35.77)					
Observations	2651	1152		942	900	849	

Notes: This table shows means and standard deviations of risk attitude and village characteristics. Column (3) shows p-values for T-Tests between Column (1) and Column (2). Column (7) shows p-values for T-Tests between columns (6) and (4).

Table 2: Summary Statistics for Full and Repeated Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Participation in study			Sample in Different Years			
	(Full Sample)	(Repeated Sample)	(T-Test)	(2008)	(2010)	(2013)	(T-Test)
Panel A: Subjective Well-being Indicators							
Ill-being (Past)	2.39	2.43	0.07	2.61	2.30	2.23	0.00
	(0.85)	(0.86)		(0.92)	(0.76)	(0.80)	
Pessimism (Future)	2.62	2.67	0.03	2.96	2.55	2.30	0.00
	(1.01)	(1.04)		(1.09)	(0.91)	(0.88)	
Panel B: Microeconomic Shocks							
No of shocks	1.50	1.46	0.27	1.69	1.60	1.18	0.00
	(1.50)	(1.51)		(1.59)	(1.57)	(1.25)	
No of high impact shocks	0.66	0.67	0.40	0.80	0.54	0.62	0.00
	(0.94)	(0.92)		(1.01)	(0.86)	(0.91)	
No of medium impact shocks	0.49	0.48	0.48	0.59	0.42	0.46	0.00
	(0.80)	(0.77)		(0.88)	(0.73)	(0.78)	
No of low impact shocks	0.09	0.08	0.02	0.12	0.08	0.07	0.00
	(0.33)	(0.28)		(0.38)	(0.31)	(0.29)	
No of shocks (demographic)	0.34	0.31	0.02	0.44	0.25	0.34	0.00
	(0.61)	(0.59)		(0.74)	(0.47)	(0.56)	
No of shocks (social)	0.26	0.26	0.98	0.26	0.21	0.31	0.08
	(0.52)	(0.53)		(0.54)	(0.48)	(0.55)	
No of shocks (agricultural)	0.50	0.49	0.55	0.64	0.45	0.41	0.00
	(0.79)	(0.77)		(0.91)	(0.70)	(0.69)	
No of shocks (economic)	0.22	0.21	0.25	0.32	0.17	0.15	0.00
	(0.49)	(0.47)		(0.57)	(0.42)	(0.43)	
No of shocks (covariate)	0.62	0.59	0.14	0.83	0.52	0.49	0.00
	(0.94)	(0.89)		(1.09)	(0.85)	(0.80)	
No of shocks (idiosyncratic)	0.67	0.64	0.11	0.82	0.50	0.69	0.01
	(0.89)	(0.89)		(1.04)	(0.71)	(0.86)	
Observations	2651	1152		942	900	849	

Notes: This Table shows means of subjective assessment of well-being and self-reported micro-economic shocks. Standard deviations are reported in parenthesis. Column (3) shows p-values for t-tests for equal means between Column (1) and Column (2). Column (7) shows p-values for t-tests for equal means between columns (6) and (4).

Table 3: Stability of Preferences Over Time

	(1) CRRA 2013	(2) CRRA 2013	(3) CRRA 2013
CRRA 2008	0.147***	0.138***	0.137**
	(0.043)	(0.046)	(0.066)
Female		0.094**	-0.015
		(0.046)	(0.069)
Age		0.004**	0.005
		(0.002)	(0.003)
Height (cm)		-0.000	-0.000
		(0.002)	(0.005)
Farmer		-0.019	0.050
		(0.048)	(0.070)
Self-employed		-0.165*	0.249**
		(0.085)	(0.115)
Public Servant		-0.596***	-0.411*
		(0.130)	(0.215)
Martial Status		-0.150***	-0.012
		(0.049)	(0.075)
Years of Schooling		-0.009	-0.010
		(0.008)	(0.011)
Household Size		-0.028**	-0.024
		(0.012)	(0.021)
Health Status		0.015	0.008
		(0.028)	(0.038)
Log per capita consumption		-0.071**	-0.042
		(0.034)	(0.052)
Constant	0.040	0.612	0.392
	(0.029)	(0.427)	(0.875)
Selection Equation	No	No	Yes
Estimation	Pooled OLS	Pooled OLS	Heckmann
R-Squared	0.01	0.07	0.06
Observations	471	471	471

Notes: This table reports coefficients of OLS and Heckman 2-Step regressions. Robust standard errors are in parentheses. All socio-economic variables are 2008 values. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table 4: Selection Bias

	(1) No Shock	(2) Has a Shock	(3) T-Test
Panel A: Village Characteristics			
No of village shocks	1.43	1.50	0.16
	(0.66)	(0.76)	
No of seasonal workers	1.96	2.11	0.54
	(3.05)	(4.06)	
Major problems	5.05	5.70	0.24
	(10.93)	(13.45)	
No of social cultural activities	6.60	6.85	0.38
	(6.29)	(6.35)	
Public Water Provision	1.08	1.07	0.24
	(0.27)	(0.25)	
Waste Water	1.16	1.16	0.99
	(0.65)	(0.65)	
Has nursery	1.49	1.51	0.48
	(0.50)	(0.50)	
Has bank	1.91	1.91	0.27
	(0.28)	(0.26)	
Panel B: Individual Characteristics			
CRRA	0.30	0.35	0.04
	(0.61)	(0.56)	
Female	0.55	0.60	0.02
	(0.49)	(0.49)	
Age	53.33	51.92	0.01
	(13.86)	(13.11)	
Marital Status	0.83	0.82	0.61
	(0.37)	(0.38)	
Years of Schooling	5.73	5.59	0.28
	(3.16)	(2.98)	
Log per capita wealth	9.01	9.02	0.89
	(1.13)	(1.03)	
Log per capita consumption	7.49	7.52	0.32
	(0.63)	(0.63)	
Household Size	3.91	4.16	0.00
	(1.67)	(1.77)	
Health Status	1.49	1.59	0.00
	(0.66)	(0.72)	

Notes: This table reports means and standard deviations. Has a shocks is a dummy that is 1 if the respondent reports to have had a shock in the past 5 years and 0 other wise. Difference shows the t-test results between the variables.

Table 5: Determinants of Changes in Risk Preferences

	(1) CRRA	(2) CRRA	(3) CRRA
Age	-0.052***	0.002	0.002
	(0.011)	(0.043)	(0.043)
Marital Status	0.172	0.174*	0.167
	(0.105)	(0.105)	(0.108)
Years of Schooling	0.123	0.127	0.130
	(0.182)	(0.181)	(0.182)
Household Size	-0.026	-0.024	-0.023
	(0.022)	(0.022)	(0.022)
Health Status	-0.030	-0.031	-0.031
	(0.037)	(0.037)	(0.037)
Log per capita consumption	-0.039	-0.024	-0.025
	(0.049)	(0.053)	(0.053)
No of shocks	0.025	0.026*	0.026*
	(0.015)	(0.015)	(0.015)
Pessimism (Future)	0.060**	0.058**	0.056**
	(0.027)	(0.027)	(0.027)
Ill-being (Past)	0.058***	0.055***	0.058***
	(0.021)	(0.020)	(0.020)
Year Dummy 2010		-0.146	-0.146
		(0.092)	(0.092)
Year Dummy 2013		-0.287	-0.285
		(0.216)	(0.216)
Constant	2.434**	-0.477	-0.529
	(1.193)	(2.501)	(2.505)
Within R-Squared	0.126	0.128	0.132
District Dummy	No	No	Yes
Number of observations	384	384	384

Notes: This table reports within-changes in levels of risk aversion using a selected set of individual characteristics, subjective assessment of well-beings and shocks. We include year and district fixed effects. Standard errors are in parenthesis and clustered on a village level. Female and height are dropped out. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table 6: Determinants of Changes in Risk Preferences

	(1) CRRA	(2) CRRA	(3) CRRA
Age	-0.008	0.003	-0.002
	(0.043)	(0.044)	(0.044)
Marital Status	0.148	0.166	0.164
	(0.103)	(0.106)	(0.108)
Years of Schooling	0.099	0.123	0.117
	(0.179)	(0.170)	(0.183)
Household Size	-0.026	-0.025	-0.023
	(0.022)	(0.022)	(0.023)
Health Status	-0.032	-0.038	-0.032
	(0.036)	(0.035)	(0.037)
Log per capita consumption	-0.030	-0.015	-0.021
	(0.053)	(0.053)	(0.052)
Pessimism (Future)	0.056**	0.057**	0.056**
	(0.027)	(0.027)	(0.027)
Ill-Being (Past)	0.058***	0.051**	0.058***
	(0.020)	(0.021)	(0.020)
No of demographic shocks	-0.003		
	(0.031)		
No of social shocks	0.024		
	(0.042)		
No of agricultural shocks	0.035		
	(0.026)		
No of economic shocks	0.131***		
	(0.044)		
High impact shocks		0.083***	
		(0.023)	
Medium impact shocks		0.028	
		(0.024)	
Low impact shocks		-0.006	
		(0.069)	
No of covariate shocks			0.025
			(0.018)
No of idiosyncratic shocks			0.020
			(0.021)
Year Dummy 2010	-0.095	-0.117	-0.123
	(0.093)	(0.095)	(0.094)
Year Dummy 2013	-0.217	-0.280	-0.264
	(0.214)	(0.224)	(0.219)
Constant	0.201	-0.628	-0.240
	(2.442)	(2.510)	(2.524)
Within R-Squared	0.177	0.178	0.174
District Dummy	No	No	Yes
Number of observations	384	384	384

Notes: This table reports within-changes in levels of risk aversion using a selected set of individual characteristics, subjective assessment of well-beings and shocks. We include year and district fixed effects. Standard errors are in parentheses and clustered on a village level. Female and height are dropped out. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table 7: Determinants of Changes in Risk Preferences (Income Effects)

	(1) CRRA	(2) CRRA	(3) CRRA
Age	-0.027	0.016	0.018
	(0.018)	(0.027)	(0.027)
Marital Status	0.139	0.143	0.104
	(0.143)	(0.145)	(0.144)
Years of Schooling	-0.365***	-0.368***	-0.372***
	(0.122)	(0.123)	(0.122)
Household Size	0.013	0.020	0.022
	(0.028)	(0.027)	(0.027)
Health Status	-0.044	-0.038	-0.040
	(0.046)	(0.048)	(0.047)
Log per capita consumption	-0.099*	-0.071	-0.069
	(0.056)	(0.064)	(0.064)
Total Financial Aid	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)
Income Loss due to Shocks	0.006	0.006	0.006
	(0.004)	(0.004)	(0.004)
No of economic shocks	0.144***	0.130**	0.125**
	(0.054)	(0.058)	(0.057)
Pessimism (Future)	0.043	0.039	0.033
	(0.041)	(0.041)	(0.041)
Ill-Being (Past)	0.069*	0.066*	0.073**
	(0.036)	(0.036)	(0.035)
Year Dummy 2010		-0.144**	-0.142**
		(0.066)	(0.067)
Year Dummy 2013		-0.240*	-0.243*
		(0.139)	(0.141)
Constant	3.965***	1.545	1.448
	(1.359)	(1.712)	(1.709)
Within R-Squared	0.134	0.140	0.154
District Dummy	No	No	Yes
Number of observations	384	384	384

Notes: This table reports within-changes in levels of risk aversion using a selected set of individual characteristics, subjective assessment of well-beings and shocks. We include year and district fixed effects. Standard errors are in parenthesis and clustered on a village level. Female and height are dropped out. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table 8.A: Impact of Shocks on Risk Preferences (Subpopulation)

	(1) Low Income CRRA	(2) High Income CRRA
Panel A: Income		
No of demographic shocks	0.054	-0.091
	(0.047)	(0.059)
No of social shocks	-0.036	-0.076
	(0.072)	(0.064)
No of agricultural shocks	0.030	0.041
	(0.043)	(0.042)
No of economic shocks	0.179**	0.217***
	(0.071)	(0.069)
Constant	3.840*	3.404
	(-2.174)	(-7.781)
Controls	Yes	Yes
Within R-Squared	0.200	0.172
Year Dummy	Yes	Yes
District Dummy	Yes	Yes
Number of observations	285	299

Table 8.B: Impact of Shocks on Risk Preferences (Subpopulation)

	(1) Low Insurance CRRA	(2) High Insurance CRRA
Panel A: Insurance		
No of demographic shocks	0.032	-0.074
	(0.067)	(0.056)
No of social shocks	0.029	-0.031
	(0.079)	(0.061)
No of agricultural shocks	0.080*	-0.004
	(0.049)	(0.036)
No of economic shocks	0.182**	0.174**
	(0.088)	(0.081)
Constant	1.690	-4.325
	(-2.231)	(-13.221)
Controls	Yes	Yes
Within R-Squared	0.142	0.182
Year Dummy	Yes	Yes
District Dummy	Yes	Yes
Number of observations	327	301

Notes: Table 8.A and 8.B report within-changes in levels of risk aversion using demographic, social, agricultural, and demographic shocks. We include year and district fixed effects. Robust standard errors are in parentheses. Controls include Age, Height, Years of Schooling, Household Size, Marital Status and Log per capital consumption. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Figure 3: This figure illustrates the frequency distribution of the qualitative measure of risk aversion in 2008, 2010 and 2013 for repeated respondents. The qualitative indicator tries to elicit the risk attitude by asking them: "Are you generally a person who is fully prepared to take risks or do you try to avoid taking risk? " Please choose a number on a scale from 0 (unwilling to take risk) or 10 (fully prepared to take risk).

Table 9: Determinants of Changes in Risk Preferences (WTR)

	(1) WTR	(2) WTR	(3) WTR
Age	0.422***	0.276***	0.272***
	(0.052)	(0.085)	(0.085)
Marital Status	-0.406	-0.398	-0.227
	(0.641)	(0.640)	(0.595)
Years of Schooling	2.004	2.086	2.077
	(1.299)	(1.342)	(1.344)
Household Size	0.061	0.092	0.085
	(0.138)	(0.139)	(0.140)
Health Status	0.241	0.194	0.198
	(0.189)	(0.186)	(0.186)
Log per capita consumption	-0.302	-0.080	-0.065
	(0.262)	(0.283)	(0.283)
No of shocks	-0.155*	-0.145*	-0.143*
	(0.084)	(0.082)	(0.082)
Pessimism (Future)	-0.001	-0.020	-0.013
	(0.137)	(0.138)	(0.138)
Ill-Being (Past)	-0.253**	-0.290***	-0.310***
	(0.109)	(0.108)	(0.110)
Year Dummy 2010		-0.203	-0.211
		(0.260)	(0.261)
Year Dummy 2013		0.636	0.643
		(0.451)	(0.452)
Constant	-25.269***	-19.540**	-19.851**
	(7.697)	(9.601)	(9.639)
Within R-Squared	0.164	0.171	0.176
District Dummy	No	No	Yes
Number of observations	384	384	384

Notes: This table reports within-changes in levels of risk aversion using a selected set of individual characteristics, subjective assessment of well-beings and shocks. We include year and district fixed effects. Standard errors are in parenthesis and clustered on a village level. Female and height are dropped out. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table 10: Determinants of Changes in Risk Preferences (WTR)

	(1) WTR	(2) WTR	(3) WTR
Age	0.304***	0.266***	0.294***
	(0.086)	(0.081)	(0.083)
Marital Status	-0.113	-0.173	-0.191
	(0.596)	(0.603)	(0.601)
Years of Schooling	2.205	2.182	2.163
	(-1.329)	(-1.373)	(-1.338)
Household Size	0.093	0.091	0.083
	(0.141)	(0.140)	(0.140)
Health Status	0.191	0.215	0.192
	(0.187)	(0.182)	(0.186)
Log per capita consumption	-0.064	-0.112	-0.086
	(0.284)	(0.282)	(0.283)
Pessimism (Future)	-0.017	-0.018	-0.014
	(0.137)	(0.136)	(0.138)
Ill-Being (Past)	-0.313***	-0.285**	-0.311***
	(0.112)	(0.111)	(0.110)
No of demographic shocks	0.021		
	(0.178)		
No of social shocks	0.088		
	(0.216)		
No of agricultural shocks	-0.048		
	(0.158)		
No of economic shocks	-0.415*		
	(0.233)		
High impact shocks		-0.252**	
		(0.126)	
Medium impact shocks		0.073	
		(0.172)	
Low impact shocks		0.085	
		(0.402)	
No of covariate shocks			-0.102
			(0.145)
No of idiosyncratic shocks			-0.037
			(0.128)
Year Dummy 2010	-0.337	-0.245	-0.288
	(0.268)	(0.264)	(0.263)
Year Dummy 2013	0.471	0.726	0.576
	(0.457)	(0.445)	(0.443)
Constant	-22.415**	-20.004**	-21.401**
	(9.417)	(9.598)	(9.425)
Within R-Squared	0.177	0.178	0.174
District Dummy	Yes	Yes	Yes
Number of observations	384	384	384

Notes: This table reports within-changes in levels of risk aversion using a selected set of individual characteristics, subjective assessment of well-beings and shocks. We include year and district fixed effects. Standard errors are in parenthesis and clustered on a village level. Female and height are dropped out. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table 11: Determinants of Changes in Risk Preferences (SWR)

	(1) SWR	(2) SWR	(3) SWR
Age	0.045	-0.108	-0.116
	(0.128)	(0.464)	(0.464)
Marital Status	-1.979*	-2.057*	-1.762
	(1.122)	(1.123)	(1.121)
Years of Schooling	1.853	1.560	1.543
	(2.185)	(1.985)	(1.990)
Household Size	0.196	0.067	0.053
	(0.268)	(0.262)	(0.262)
Health Status	0.024	0.168	0.173
	(0.431)	(0.431)	(0.431)
Log per capita consumption	0.932	0.076	0.102
	(0.581)	(0.606)	(0.608)
No of shocks	-0.410**	-0.449**	-0.445**
	(0.179)	(0.175)	(0.176)
Pessimism (Future)	-0.945***	-0.860***	-0.848***
	(0.315)	(0.321)	(0.322)
Ill-Being (Past)	-0.789***	-0.650**	-0.686***
	(0.259)	(0.251)	(0.254)
Year Dummy 2010		2.278**	2.264**
		(0.994)	(0.995)
Year Dummy 2013		1.252	1.264
		(-2.280)	(-2.280)
Constant	-4.953	9.949	9.422
	(15.052)	(27.116)	(27.124)
Within R-Squared	0.066	0.088	0.092
District Dummy	No	No	Yes
Number of observations	384	384	384

Notes: This table reports within-changes in levels of risk aversion using a selected set of individual characteristics, subjective assessment of well-beings and shocks. We include year and district fixed effects. Standard errors are in parenthesis and clustered on a village level. Female and height are dropped out. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table 12: Determinants of Changes in Risk Preferences (SWR)

	(1) SWR	(2) SWR	(3) SWR
Age	0.053	-0.131	-0.035
	(0.455)	(0.470)	(0.473)
Marital Status	-1.553	-1.728	-1.733
	(1.056)	(1.100)	(1.124)
Years of Schooling	2.027	1.689	1.741
	(1.724)	(1.493)	(2.004)
Household Size	0.096	0.080	0.036
	(0.257)	(0.254)	(0.262)
Health Status	0.206	0.280	0.195
	(0.420)	(0.416)	(0.431)
Log per capita consumption	0.170	-0.064	0.052
	(0.596)	(0.599)	(0.604)
Pessimism (Future)	-0.857***	-0.875***	-0.856***
	(0.315)	(0.314)	(0.324)
Ill-Being (Past)	-0.680***	-0.572**	-0.682***
	(0.253)	(0.255)	(0.255)
No of demographic shocks	-0.146		
	(0.330)		
No of social shocks	-0.416		
	(0.518)		
No of agricultural shocks	-0.752**		
	(0.289)		
No of economic shocks	-1.967***		
	(0.527)		
High impact shocks		-1.329***	
		(0.252)	
Medium impact shocks		-0.387	
		(0.266)	
Low impact shocks		0.125	
		(0.732)	
No of covariate shocks			-0.426*
			(0.216)
No of idiosyncratic shocks			-0.421
			(0.260)
Year Dummy 2010	1.396	1.832*	1.843*
	(1.021)	(1.023)	(1.028)
Year Dummy 2013	0.139	1.230	0.890
	(2.262)	(2.338)	(2.335)
Constant	-2.193	10.747	4.551
	(25.887)	(26.124)	(27.419)
Within R-Squared	0.120	0.119	0.092
District Dummy	Yes	Yes	Yes
Number of observations	384	384	384

Notes: This table reports within-changes in levels of risk aversion using a selected set of individual characteristics, subjective assessment of well-beings and shocks. We include year and district fixed effects. Standard errors are in parenthesis and clustered on a village level. Female and height are dropped out. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table 13: Determinants of Changes in Risk Preferences (Upper Bound CRRA)

	(1) Upper Bound CRRA	(2) Upper Bound CRRA	(3) Upper Bound CRRA
Age	-0.016	-0.007	-0.012
	(0.027)	(0.027)	(0.027)
Marital Status	0.106	0.121	0.122
	(0.084)	(0.085)	(0.087)
Years of Schooling	0.083	0.097	0.095
	(0.162)	(0.157)	(0.162)
Household Size	-0.016	-0.015	-0.013
	(0.018)	(0.017)	(0.018)
Health Status	-0.036	-0.042	-0.036
	(0.027)	(0.026)	(0.027)
Log per capita consumption	-0.020	-0.009	-0.012
	(0.040)	(0.041)	(0.040)
Pessimism (Future)	0.042*	0.043**	0.042*
	(0.021)	(0.021)	(0.022)
Ill-Being (Past)	0.044***	0.038**	0.044***
	(0.015)	(0.015)	(0.015)
No of demographic shocks	-0.004		
	(0.023)		
No of social shocks	0.015		
	(0.032)		
No of agricultural shocks	0.031		
	(0.020)		
No of economic shocks	0.098***		
	(0.031)		
High impact shocks		0.067***	
		(0.017)	
Medium impact shocks		0.009	
		(0.018)	
Low impact shocks		0.018	
		(0.051)	
No of covariate shocks			0.022
			(0.014)
No of idiosyncratic shocks			0.011
			(0.017)
Year Dummy 2010	-0.086	-0.102	-0.106*
	(0.061)	(0.061)	(0.061)
Year Dummy 2013	-0.133	-0.183	-0.164
	(0.137)	(0.141)	(0.138)
Constant	0.721	0.143	0.433
	(1.647)	(1.671)	(1.680)
Within R-Squared	0.162	0.164	0.148
District Dummy	Yes	Yes	Yes
Number of observations	369	369	369

Notes: This table reports within-changes in levels of risk aversion (upper bound only) using a selected set of individual characteristics, subjective assessment of well-beings and shocks. We include year and district fixed effects. Standard errors are in parenthesis and clustered on a village level. Female and height are dropped out. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Appendix A Details of Risk Elicitation Tasks

In our data collection process for the experiment, we tried to keep the enumerator instructions as short and simple as possible in order to facilitate the understanding.

1. **Self-reported risk attitude**: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risk? (Please choose a number on a scale from 0 to 10).

Table A1: Self-Reported Risk Attitude

0=avoid risk | 10=seek risk

2. Certainty equivalent experiment: This is game 1. It has 20 rows. In each row a decision has to be made. In each row we would like you to choose option A or option B. Option A is a certain amount of THB. It starts with 0 and goes up by 10 THB in every row. Option B is a lottery where a coin is thrown. If 'King' falls you win 300 Baht. If 'Palace' falls you get nothing. *(Enumerator shows the coin)*. Please make your choice of Option A or B for each row. If this game is selected to be played with real money, you will be asked to draw a number from a bag. The bag contains the numbers 1 to 20 for the 20 rows. We will play with real money according to your choice. For example: If you draw the number X (Enumerator ID) from the bag, we play the game at this row for money. That means: If you chose option A you will receive (THB). If you chose option B we will toss a coin. If 'King' you win 300 Baht. If 'Palace' you win nothing.

Table A2: Certainty Equivalent Task

Row	**Option A**	**Tick Box**	**Tick Box**	**Option B**
1	0			300 : 0
2	10			300 : 0
3	20			300 : 0
4	30			300 : 0
5	40			300 : 0
6	50			300 : 0
7	60			300 : 0
8	70			300 : 0
9	80			300 : 0
10	90			300 : 0
11	100			300 : 0
12	110			300 : 0
13	120			300 : 0
14	130			300 : 0
15	140			300 : 0
16	150			300 : 0
17	160			300 : 0
18	170			300 : 0
19	180			300 : 0
20	190			300 : 0

Appendix B Description of Variables

B.1 Individual and Household Characteristics

Female is a dummy variable. It takes the value 1 for female and 0 for male.

Age is the respondents' age in years.

Height is respondents' height in cm.

Years of Schooling is denoted as years in eduation.

Marital Status is a dummy variable. It takes the value 1 if married and 0 otherwise.

Household Size is the headcount of persons who are living in the household for at least 180 days.

Health Status asks the question: How healthy do you feel? 1=feel good; 2 =manageable; 3=sick.

Number of Insurances is the total sum of voluntary insurances of the household.

LPCC is the log per capita consumption and it refers to the natural logarithm of household consumption per day divided by OECD adult equivalents AE (AE = 1+0.7*(adults-1) + 0.5*children).

Farmer is a dummy variable. It takes the value 1 for being a farmer and 0 otherwise.

Self-employed is a dummy variable. It takes the value 1 for being self-employed and 0 otherwise.

B.2 Shocks and Well-being Measures

Ill-being (Past) asks the question: Do you think your household is better off than 5 years ago? 1= Much better off; 2=Better off; 3=Same; 4=Worse off; 5=Much worse off

Pessimism (Future) asks the question: Do you think your household will be better off in 5 years? 1=Much better off; 2=Better off; 3=Same; 4=Worse off; 5=Much worse off.

Number of demographic shocks is the sum of these self-reported experienced shocks:

Household member left the household

Illness of household member

Person joined the household

Number of social shocks is the sum of these self-reported experienced shocks:

Accident

Conflict with neighbours in the village

Household was cheated

Household damage

Law suit

Money spent for ceremony in the household

Relatives/friends stopped sending money

Theft

Number of agricultural shocks is the sum of these self-reported experienced shocks:

Crop pests

Drought

Flooding of agricultural land

Landslide, erosion

Livestock disease

Snow/ice rain

Storage pests, incl. rats

Storm

Unusually heavy rainfall

Number idiosyncratic shocks is the sum of these self-reported experienced

shocks which affected the household only or some of the households in the village:
Accident
Collapse of business
Conflict of neighbour
Death of household member
Illness of household member
Strong increase of interest rate on loans
Household was cheated
Household Damage
Household member left the household
Illness of household member
Job loss (agricultural)
Job loss (non-agricultural)
Landslide, Erosion
Law suit
Livestock Disease
Money spent for ceremony in the household
Person joined the household
Relatives/Friends stopped sending money
Supporting others
Theft
Unable to pay back loan

Number of covariate shocks is the sum of these self-reported experienced shocks which affected the entire village or district:
Snow / ice rain
Storage pests (including rats)
Storm
Strong decrease of prices for Output

Strong increase of prices for Input

Unusually heavy rainfall

Flooding of agricultural land

Drought

Crop pests

Change in market regulations

B.3 Village Characteristics

No of village shocks is the reported number of shocks a village experienced by the village head.

No of seasonal workers is the number of seasonal workers.

Major problems asks the question: Are there major problems in the village? 1=violence; 2=drug abuse/trafficking; 3=human trafficking;4=epidemics; 5=politics

No of social cultural activities asks the question: What are the major social and cultural agricultural activities in the village?

Community meeting

Village radio system

Sport event, specify type of sport

Traditional festivals, specify festival

Festival related to Buddhist tradition

Public Water Provision asks the question: Is there a Public water supply available? Dummy for 1 is if the village has public water provision and 0 otherwise.

Waste Water asks the question: What is the main kind of waste water disposal? 1=discharge to the ground; 2=discharge to pond; 3=waste water pipes.

Solid waste management asks the question: What is the main kind of solid waste ? 1= burn; 2=dumping site; 3=public disposal.

Has nursery is a dummy for 1 if the village has a nursery and 0 otherwise.

Has bank is a dummy for 1 if the village has a bank and 0 otherwise.

Distance to district town is the travel distance from the district town in km.

Distance to provincial capital is the travel distance from the provincial capital in km.

Appendix C Validity of Risk Measures

One concern of self-reported measures based on a non-incentivized hypothetical question is whether they actually reflect an individual's underlying risk traits. Several studies have documented that risk measures obtained by hypothetical survey questions are reliable predictors of actual risk-taking behaviors (e.g., Barsky et al., 1997; Donkers et al., 2001; Dohmen et al., 2011; Anderson and Mellor, 2011). In order to check the validity of our risk measures, described in detail above, we simply run the regression of gambling on our risk measures. Table C.1 confirms the validity of our risk measures. We examine whether the general risk attitude and certainty equivalent task can explain revealed behavior in the year 2008, i.e. whether a household member buys lottery tickets. We see for 2008 that all risk measures show a significant correlation between our risk measures and gambling with expected signs.

Table A3: Determinants of Gambling

	(1) Lottery	(2) Lottery
Survey Item	0.065***	
	(0.019)	
Certainty Equivalent		0.021**
		(0.009)
Female	-0.170	-0.175
	(0.106)	(0.108)
Age	-0.009**	-0.011***
	(0.004)	(0.004)
Height	-0.005	-0.004
	(0.006)	(0.006)
Marital Status	0.066	0.071
	(0.146)	(0.148)
Years of Schooling	0.009	0.011
	(0.017)	(0.017)
Household Size	0.047	0.046
	(0.030)	(0.030)
Health Status	0.026	0.013
	(0.057)	(0.057)
Log per capita consumption	0.289***	0.297***
	(0.077)	(0.078)
R-Squared	0.05	0.04
Number of observations	807	807

Notes: The dependent variable, purchase of lottery tickets, takes the value of 1 if the household reports expenditures for lottery tickets and zero otherwise. We use a probit measure with marginal effects at the mean (0.50). Standard errors are in parentheses and clustered on a village level, we only use as subpopulation of age between 18 and 80. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Chapter 2

Estimating Risky Behavior with Multiple-Item Risk Measures: Experimental Evidence

[1]This chapter is based on the article by Menkhoff and Sakha (2014).We would like to thank participants at several conferences and seminars for valuable comments and suggestions, in particular Veronika Bertram-Huemmer, Elisa Cavatorta, Antonia Grohmann, Glenn Harrison, Stephan Klasen, Ulrich Schmidt, Ferdinand Vieider and Juliane Zenker. Financial support by the German Research Foundation (DFG, RTG 1723) is gratefully acknowledged.

2.1 Introduction

Economic decision-making and interactions involve risk. Thus, a substantial body of research tries to understand how decision makers consider risk in their choices. These analyses require a reliable measure of individual risk attitude. So far, most studies choose one of the existing methods in eliciting individual risk attitude and take the response to this specific item as "the" individual risk attitude. However, a troubling result derived from the experimental literature indicates that the degree of risk aversion varies for the same person across elicitation techniques (e.g. Isaac and James (2000); Berg et al. (2005); Anderson and Mellor (2008); Loomes and Pogrebna (2014)).

One reason for inconsistency in revealing risk attitude is related to domain specificity in risk-taking behavior where risk measures perform differently across various domains (Weber et al. (2002); Dohmen et al. (2011)). Further evidence shows that the existence of noise in survey items or experimental measures may distort findings (Harless and Camerer (1994); Loomes et al. (2002)) and affect the prediction of predicting economic outcomes (Cutler and Glaeser (2005); Beauchamp et al. (2012)). It is, thus, assumed that the choice of risk measure will make a difference. So how can researchers deal with this inconsistency?

In many cases it seems almost impossible to decide ex ante which specific risk measure is most appropriate or whether one risk measure is superior to another in eliciting risk attitude. This is the starting point of our research. We aim to investigate which risk measure may be more broadly applicable to many dimensions of risky behavior when compared to another; or whether combinations of risk measures, and if so, which kind of combinations are most useful in a situation of ignorance. This purpose is not only of academic interest but also highly relevant for various kinds of empirical applications in different contexts.

The benchmark for deciding on the effectiveness of a risk measure is its power in explaining risk-related behavior in many different dimensions (i.e. financial, gambling, health etc.) while controlling for individual socio-demographic characteristics (see Dohmen et al. (2011)). We build on this line of literature which includes, for example, Barsky et al. (1997), Tanaka et al. (2010), and Dohmen et al. (2011). However, as we assume a state of ex ante ignorance, it is crucial to extend the scope of risk measures and risk behaviors beyond the often specific coverage in earlier work. Our study with 760 individuals compares seven established measures of risk attitude and their ability to explain eleven kinds of risky behavior within five behavioral dimensions.

The obtained results demonstrate the limitations of narrow approaches where isolated relations between risk measure and behavior may be depending on the specific risk measure chosen. While all single-item risk measures have some power in explaining risky behavior in our sample, there is enormous heterogeneity. Some measures perform better than others, in some cases we find domain specificity, however, and this is our core argument, it is largely unclear which risk measure to choose ex ante. As a consequence of these sensitivity analyses, we propose multiple-item risk measures: averaging across single-item risk measures, i.e. creating a 'multiple-item risk measure', improves the predictive power in explaining behavior substantially.

Our sample consists of 760 individuals in rural Thailand. We employ seven diverse but well-established methods in eliciting risk preferences using incentivized risk tasks; the certainty equivalent task (CEquiv), two Eckel-Grossmann choice tasks (EG), one with loss treatment and one without (Eckel and Grossman, 2002, 2008), and an investment choice task similar to Gneezy and Potters (1997). In addition to these four experiment-based risk measures, we employ three non-incentivized survey items of risk attitude from Dohmen et al. (2011), i.e. general willingness to take risk (WTR Gen), the willingness to take risk in financial affairs (WTR Fin) and a hypothetical investment question (HInvQ). We investigate the correlation between our risk elicitation tasks at the individual level. Measures of risk attitude are positively correlated to each other but most of them to a low degree, indicating remarkable differences between measures and thus confirming the results found in the literature (e.g. Deck et al. (2009); Crosetto and Filippin (2013)).

Consequentially, relating risk measures to various dimensions of risky behavior leads to differences in their explanatory power which can partly be related to domain specificity. The hypothetical investment question, for instance, significantly explains behavior in the financial domain. In most cases, however, it remains unclear why some measures do explain behavior well and others do not.

Thus, in the next step, we average across (standardized) risk measures. We find that a simple average across the seven measures has the highest predictive power in our sample; it significantly explains 9 out of 11 kinds of risky behavior, whereas the single-item measures explain on average 3 kinds of behavior, within a range of 2 to 6.

In order to reduce the necessary input of eliciting various measures of risk attitude and thus to come towards a more practical approach, we employ a factor analysis. We find that three factors extract most of the information from the seven single-item risk measures. Each of the three factors is dominated by one risk item. Thus, we combine one risk elicitation method each, from the resulting factors (i.e. the hypothetical investment question, general willingness to take risk item, and the Eckel-Grossman task without loss). We can see that creating a 'three-item measure' and relating it to risky behavior, still provides us with high predictive power. This result holds tentatively even if we restrict the average to the two most relevant items.

This motivates the combination of any two risk-items with each other. Our results suggest that on average, the combination of two items is able to explain more dimensions of risky behavior (on average 4 dimensions) than merely employing a single-item measure (explaining 3 dimensions on average). However, predictive power varies considerably between any two-item risk measures. Thus, it is crucial to formulate principles for building successful multiple-item risk measures.

In this sense, we further analyze the effect of combining risk items with different framings, different domains, and repetition to find out which of these characteristics may be important for creating a risk measure which robustly explains behavior. We find that including items with different framings contributes to a more reliable and predictive multiple-item risk measure. In contrast, the combination of items from different domains or averaging across repeated answers seems less helpful in improving predictive power.

We are aware that our results are specific, i.e. the preference for specific multiple-item risk measures may be the consequence of our selection of risk measures, risky behavior, and sample population. Therefore, we hesitate to recommend the inclusion of any specific risk item. However, one concrete conclusion from our result is that researchers who try to reveal risk attitude should consider using several, such as two or three, risk items with different framings to enhance external validity.

Our research is related to three strands of literature: (1) the wealth of studies examining the consistency of risk measures (e.g. Deck et al. (2009); Hey et al. (2009); De Brauw and Eozenou (2014)). Loomes and Pogrebna (2014), for instance, investigate the within-procedural robustness of three risk elicitation tasks when used to elicit risk attitudes. They find degrees of variability and disparity that are difficult to explain within the terms of any expected utility model. (2) Studies assessing the validity of risk measures by predicting risky behavior (Barsky et al. (1997); Dohmen et al. (2011); Sutter et al. (2013); Vieider et al. (2013)) and (3) studies addressing the low explanatory power by explicitly estimating the information from noise (Andersen et al. (2008b); Kimball et al. (2008)).

Our paper is organized in eight sections: Section 2.2 presents the survey data and risk elicitation methods. Section 2.3 displays descriptive statistics of our sample and experiments correlations between risk measures. Section 2.4 exhibits results on predictive ability of single-item risk measures. Section 2.5 outlines the performance of the various multiple-item risk measures. Section 2.6 concludes.

2.2 Descriptions of the Survey and the Risk Elicitation Tasks

Section 2.2 describes the implementation process of the survey and our risk elicitation tasks (Section 2.2.1), the content of these tasks (Section 2.2.2) and their relation to risk preferences (Section 2.2.3).

2.2.1 Implementation of the Survey and the Risk Elicitation Tasks

In August 2013 we conducted a risk survey in the Northeastern province Ubon Ratchathani of Thailand. The risk survey included experiments with households that also participated a few months earlier in the 5th wave of a larger household panel survey. The risk survey was conducted with 760 households in 98 villages. Due to the linkage to the household survey, we are able to extract wider socio-demographic characteristics of the respondents including: household demographics, education, consumption, assets, credit and investment, employment, and health behavior.

The implementation of risk elicitation tasks includes a thorough training of enumerators, double translation of the survey from Thai to English and vice versa and pretests of these tasks in three villages, thus ensuring the understanding of tasks for the participants. The tasks are conducted individually and always in the same order. The duration of the interview is half an hour, tasks are partly incentivized and the average payoff is about 150 Thai Baht (THB), which is slightly less than a full day earning for an unskilled worker. Detailed information about the data collection process is provided in Appendices A.1 and A.2.

2.2.2 Risk Elicitation Tasks

We begin our risk survey with three non-incentivized hypothetical questions concerning risk attitude which have been used by Dohmen et al. (2011) (Details are shown in Appendix A.3). The first item is the general risk question from the German socio-economic panel study (SOEP) and states the question "Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?" Responses are given on an 11-point Likert scale where the value 0 means "unwilling to take risks" and the value 10 means "fully prepared to take risks". The second question is domain-specific and refers to the financial affairs of the household. The third non-incentivized risk measure is the hypothetical investment question. Variations of the latter question appear in several large panel surveys of U.S. households: the Health and Retirement Study (HRS), the Panel Study of Income Dynamics (PSID), and the National Longitudinal Survey of Youth 1979 (NLSY79).

Afterwards, we conducted four incentivized experiments. The first is a certainty equivalent task and has been implemented by Bruhin et al. (2010); Abdellaoui et al. (2011); Dohmen et al. (2011); Sutter et al. (2013) or Vieider et al. (2015). Table A.1 in Appendix A illustrates the basic payoff matrix presented to subjects. The first row shows that the lottery offers a 50-50% chance of receiving either 0 or 300 THB and alternatively a safe payoff of 0 THB. The expected value of this lottery is 150 THB. Therefore, it is rational to choose the lottery. The second row, however, already offers 10 THB as safe payoff. This provides the opportunity for risk-averse individuals to opt for the safe payoff over the lottery. The value of the safe payoff is increased in each row by 10 THB. The switching row from the lottery to the safe payoff designates individuals' risk attitude. We do not allow switching back and forth.

The second and third experiments are adaptations of the Eckel and Grossman (2002, 2008) tasks. Subjects choose between one out of five possible gambles. We employ two versions of the task - with and without loss treatment. All the gambles involve a 50/50 chance of a low and high payoff. The range of gambles includes a safe alternative involving a sure payoff of 50 THB with zero variance in the loss treatment and 80 THB in the no-loss treatment. From here, the gambles increase in both expected return and risk (standard deviation) moving from Gamble 1 to 5. Losses are possible up to 30 THB. In the no loss framing, all payoffs are shifted up so that the minimum payoff is zero. Risk-averse subjects would choose lower-risk, lower-return gambles; while risk-neutral subjects are expected to choose Gamble 5 which has the highest expected rate of return.

The last experiment is similar to the one introduced by Gneezy and Potters (1997). Given its framing, it can be regarded as an incentivized version of the hypothetical investment question. Instead of choosing lotteries, subjects have to decide how to allocate a given endowment of 100 THB. They can choose to invest a fraction of their endowment in a risky lottery. The risky lottery means that with a 50/50 chance, they can either lose or win three times their money invested. In all cases, the expected value of investing is higher than the expected value of not investing, thus, a risk-neutral (or risk-seeking) person should invest 100 THB, while a risk-averse person may invest less.

2.2.3 Risk Measures and Risk Preferences

All seven risk elicitation tasks aim to reveal risk preferences but the methods are quite different. As we will sketch shortly, these measures are not fully comparable to each other in a strict sense. However, all of them are widely used as risk measures and from this applied perspective we are interested in their relative performance.

Among the four incentivized experimental elicitation methods, the certainty equivalent task is our most precise measure of risk attitude. It is based on 20 choices between a risky lottery which offers either 300 Thai Baht or zero with equal probabilities and varying safe amounts s_j, $_j = 1, \ldots, 20$ ranging from zero to 190. For the sake of simplicity, we could assume for all four incentivized methods, that risk preferences are represented by a Constant Relative Risk Aversion (CRRA) utility function defined as

$$U(y) = (y^{1-r})/(1-r)$$

where y is the lottery price and $r \neq 1$ is the CRRA coefficient to be estimated. Thus, r is the coefficient of CRRA: $r = 0$ corresponds to risk-neutrality, $r < 0$ to risk-loving, and $r > 0$ to risk aversion. Thus, for the certainty equivalent task, we can calculate the range of the CRRA; that is when subjects choose to switch from the lottery to the safe payoff. For example, a subject that makes five lottery choices and then switches to the safe payoff alternatives would reveal a CRRA interval between 0.569 and 0.613.

However, our certainty equivalent task is the only measure showing a fairly complete range of preferences. The other three tasks, the GP and both EG tasks can only estimate $r \geq 0$. In other words, they cannot distinguish risk neutrality from risk seeking. Hence, we have a limitation using the CRRA because it restricts the analysis to a utility function that characterizes risk attitudes using only one parameter. Charness and Viceisza (2012) argue that this is a minor problem because risk seeking preferences are seldom observed, although given our certainty equivalent task, 21% of the subject pool is characterized by $r < 0$. Moreover, since the different tasks are also characterized by different ways of eliciting risk aversion parameters, the shape of the function linking choices to r can greatly differ across tasks (Harrison et al., 2007). The case is different regarding the three survey items because they cannot be usefully transformed to a risk aversion parameter at all.

It is for these reasons that our analysis does not apply any specific form of utility function. Instead, we focus our efforts to investigate the external validity of our risk elicitation methods in relation to real-life choices.

2.3 Descriptive Statistics

In this section, we describe individual characteristics of our sample population (Section 2.3.1), outcomes of the seven risk measures (Section 2.3.2) and correlations of risk measures (2.3.3)

2.3.1 Individual Characteristics

Table 1 presents individual characteristics of our sample in two panels. Panel A describes seven standard socio-demographic characteristics which we will regularly use as control variables in our regressions. A large proportion of our sample are women (58%), participants are on average 54 years old and 1.58 cm tall. 98% have attended school at some point with 5.7 years of school education on average. Given the average age of our respondents, 83% of our sample are married and live in households with approximately four other household members. The log per capita consumption of 7.6 reflects an annual total household consumption level of 6073$ PPP.

Panel B summarizes answers on various behavioral items and further sample characteristics (Details can be found in Appendix B). After the completion of the risk experiments, we ask the respondents six algebra questions. Our respondents have moderate algebra skills (3.63 out of six exercises are solved correctly on average). Additionally, more than 60% of the respondents describe farming as their main activity while only 8% report that they are self-employed. About 31% of our sample engage in lottery activities at least once over the past 12 months, devoting an average share of about 6% of their annual income to lottery participation (630.32$). 70% of our respondents hold insurances (i.e. 1.95 insurances on average per household). Although the majority of health care services in Thailand are delivered by the public sector, still 6% percent choose an additional health insurance package. Additionally, 8% percent choose an accident insurance. Our sample households made on average 5084$ investments in agricultural and non-agricultural sector in the last two years which is around half of the total income. It must be noted, however, that there is a large share of people who do not make any investments. Furthermore, 70% of the respondents borrowed money in the last two years. When asked for its purpose, around 8% report to have borrowed for business purposes. Moreover, at least 49% of our sample adopt precautionary measures against future shocks and risks. It ranges from having none to eight risk-mitigating activities. Finally, considering the body mass index (BMI) of our respondents, we find that the majority can be classified as having normal weight. Lastly, following the WHO definition, 29% of our sample can be classified as being overweight.

2.3.2 Results on Risk Measures

Table 2 shows the summary statistics of all seven risk elicitation methods in the order of the survey. It can be inferred that the WTR (Gen) has an average value of 6.85 for our sample. The average value is higher than in the study of Dohmen et al. (2011) for Germany and also higher when compared to responses of similar households in the same province of Thailand in 2008 (Hardeweg et al., 2013). The driving force behind this result is related to the large number of respondents choosing the highest risk seeking category of 10. This is also the case for financial matters albeit less than in the general question which corresponds to the study of Dohmen et al. (2011) where respondents are more risk-averse in the financial part than in the general question. Higher average self-reported values are in line with Charness and Viceisza (2012) who conducted the same question in rural Senegal. In the robustness section (Appendix F), we will show that our main results are not distorted due to unusually risk-seeking answers.

Turning to the next risk item, the average amount of hypothetical investment is 50,880 THB. 60% of respondents choose the median amount of 50,000 THB. All results in Table 2 concerning the hypothetical investment question are divided by 1000 to rescale estimates. Looking at the distribution, we find that the respondents tend slightly towards investing more than 50%, indicating some risk-neutral behavior. This result is similar to Hardeweg et al. (2013).

The average switching row in the CEquiv task is 7.94. This means that 76% of the respondents are risk averse as they choose to switch before row 16, while 3% are risk-neutral and 21% are risk seeking. Qualitatively similar results for risk averse behavior are found, for example, Harrison et al. (2007) in Denmark, Dohmen et al. (2011) in Germany, and Hardeweg et al. (2013) in Thailand.

For both EG tasks, we find the median lottery choice to be 3, confirming risk averse behavior. Comparing means for each pair of treatments, we are able to reject the null hypothesis of no difference by treatment (p-value=0.01). Respondents in the EG (Loss) task tend to opt in a greater scale for the risky gambles, i.e. gambles 4 and 5 in the EG (No Loss) treatment. Given the relative small loss coefficients in our EG (Loss) task, however, we can largely disregard the loss aversion parameter in our EG (Loss) task (Wakker, 2010). Hence, the overall image of risk-averse respondents is also confirmed with both EG tasks.

Concerning the last incentivized task, the GP experiment, we find that people invest less in the risky option than in the HInvQ. Only 7% choose to invest all their endowment in the risky option while nearly 30% chose to invest nothing, indicating strong risk-averse behavior (similar results are found in Charness and Villeval (2009)).

Generally, we find from our four risk experimental measures that the respondents are fairly risk-averse. The underlying evidence from the non-incentivized risk items seems to be similar.

Concerning the determinants of risk attitude (Appendix C, Table C.1), we find significant relations with the expected signs, with the exception of age and height of the respondent. We find, for instance, that marital status is positively correlated to risk aversion as well as household size and consumption, while more years of schooling seem to increase risk-tolerance.

2.3.3 Correlations of Risk Measures

Table 3 depicts Spearman rank correlations between various elicitation methods. We observe mostly positive correlations. 12 out of 21 coefficients are statistically significant and all of these coefficients are positive. Some coefficients have higher correlations, such as the two risk survey questions (0.36) and both EG tasks (0.44). Other coefficients are rather moderate, for instance between the HInvQ and the GP task (0.20). Moreover, within elicitation methods with similar framings, the degree of correlation is higher than across elicitation methods.

Low correlation between tasks is a recurrent finding in the literature. Crosetto and Filippin (2013), for instance, perform a comparative analysis of five-incentivized tasks and also find low correlations. Their explanation is that when facing multiple decisions under uncertainty, subjects could maximize their utility in every period, thereby making the best choice every time. For example, subjects could make in the first two choices risk-averse decisions and risk-loving ones in the second two. If this is the case, low correlation across tasks would be an artifact of the multiple decision framework rather than reflecting idiosyncratic features of different tasks. However, such effects should be marginal in our setting as participants were not informed about the number and kind of tasks to be played at the beginning.

Deck et al. (2009) compare four elicitation methods: Holt and Laury (HL), EG, the Balloon, and a version of 'Deal or Not Deal' TV show. They find a significant, though weak, correlation between the EG and HL and between the two visual tasks (The Ballon and Deal or Not Deal). Hey et al. (2009) compare willingness-to-pay, willingness-to accept, BDM measures and choices over pairwise lotteries. They find inconsistencies and in some cases even negative correlations between results of the different methods within individuals. This is comparable to the study of Vieider et al. (2015). They examine individual correlations between different incentivized measures and survey questions for 30 countries and find low within-country correlation coefficients between tasks – from slightly negative to 0.40. Anderson and Mellor (2008) compare results of the methods developed by Holt and Laury (2002) and survey results on gambles and find that except for a small fraction of 'consistently consistent' decision makers, the methods do not provide consistent within-individual estimates of risk attitudes.

Overall, we find significant but low correlation across tasks. Further, we hypothesize that the degree of correlation is highest when two measures have similar frames (i.e. survey vs. experimental design) or are within the same domain (i.e. here investment / financial issues).

2.4 Risk Measures and Risky Behavior

In this section, we relate all the seven measures of risk attitude to eleven risk behavior items that according to theory should be greatly affected by risk attitude. Appendix D provides a more thorough analysis. We cover five domains of risky behavior: gambling (playing lottery), risky employment, financial behavior (investment, borrowing), risk avoidance (insurance), and behavior towards health. We find unequivocal evidence that the direction of prediction is always correct, but relatively low.

2.4.1 Areas of Risky Behavior

Gambling (Playing Lottery). The relationship between gambling and risk attitude is close because participation in a lottery is a risky decision. Hence, the purchase of lottery tickets is a good indicator of risk-seeking behavior (see e.g. Clotfelter and Cook (1990)). Our survey ascertains the purchase of lottery tickets for the total household in the last 12 months. Complementing the backward looking question on past lottery expenditure, we also ask the respondent how much, he/she is willing to spend in the next lottery drawing. We estimate an OLS regression on the effect of risk attitude on lottery expenditure as a share of total household expenditure. All statistically significant coefficients are presented in columns (1) and (2) in Table 4.

Risk attitude is significantly correlated to past lottery ticket purchase for the WTR (Gen) and the EG (Loss) while employing a set of control variables. Both risk measures exhibit the expected signs indicating that risk-tolerance is associated with higher lottery expenditure. Details of the full regression results are available upon request.

With regard to the future lottery spending, we also find as before that increased future spending on lottery tickets is positively correlated with increased risk-seeking behavior in the EG (Loss) task at the 5% significance level. In addition to that, we find that also the GP task and EG (No Loss) significantly reveal the relationship between future lottery spending and risk aversion. In contrast, it seems more difficult to explain why the WTR (Gen) lost its explanatory power for future spending.

Risky Employment. Entrepreneurship is another prominent example of risky behavior since entrepreneurs and their business formation are embedded in constant financial uncertainty. In contrast, the decision of becoming a farmer is different. While they have to take decisions under uncertainty as entrepreneurs, being a farmer in a poorer rural area is often the last resort of those who are not entrepreneurial or mobile; thus we expect these persons to be less risk tolerant. Since we cannot clearly identify causality, the results are interpreted as correlates.

We implement a Probit model to estimate the correlation between risk attitude and the probability of being self-employed. Column (3) in Table 4 displays the marginal effects at the mean observation. Risk attitudes are significantly related to self-employment which is revealed by the WTR (Gen), the HInvQ, and the

GP. Thus, our results are in line with the findings of, for example, Moskowitz and Vissing-Jörgensen (2002). In column (4), we find that the CEquiv and the EG (No Loss) experiments significantly explain the decision of people becoming farmers at the 1% and 10%, respectively. This is in line with the finding of Reynaud and Couture (2012).

Financial Behavior. Even though poorer farmers do not and cannot hold a market portfolio, the more risk-averse should hold safer portfolios and thus make less risky investments. Moreover, planning an investment in the future is embedded in uncertainty about the conditions under which the planned investment may take place. We hypothesize that risk-tolerant rather than risk-averse individuals should be more prone in planning to conduct considerable investments. Third, borrowing can be generally seen as a decision which entails risk because the borrower has agreed to future repayment without knowing his future economic situation. Thus, the less risk-averse individuals should be more likely to borrow more.

As a regression model for explaining respondents' investment decision, i.e. the share of the income they invest as part of households income, we find in column (5) that the risk items which are closely correlated to investment issues - the HInvQ and the GP task - are able to explain significantly whether respondents invest a higher share of their income. Using a probit model to estimate the correlation between risk attitude and planned investment, we find in column (6) that the CEquiv task is able to explain planned investment and is statistically significant at the 1% level. We employ a probit model for borrowing and find in column (7) that the probability of borrowing is significantly explained by the WTR (Fin). Furthermore, since borrowing incorporates potential future investment, this may explain why the HInvQ is statistically significant at the 10% level in predicting the probability of borrowing.

Risk Avoidance. In the absence of formal insurance markets, in theory, risk-averse individuals may choose to implement or undertake more risk-coping mechanisms (i.e. substitute crops, diversify agricultural portfolio etc.) than a risk-seeking individual. We find a statistically significant relationship between risk attitude and the implementation of any precautionary measures against shocks and risks (see column 8); this holds for the HInvQ and the EG (No Loss) risk measures. The results are consistent with findings in the literature (see Dercon and Christiaensen, 2011).

Next, we examine the number of insurance contracts that a household holds. Using a probit regression as shown by column (9), we find that more risk-seeking consumers are less likely to buy private insurances. This holds only for the EG (No Loss) risk item.

Behavior Towards Health. Regarding health issues we consider the case where next to the free health insurance from the state, respondents also chose to have an additional health insurance with better coverage. The expectation is about the same as for the number of insurance contracts analyzed before. We also compute the BMI of our respondents and expect that higher BMI of students is strongly associated with less risk aversion (see Sutter et al., 2013). For health insurance, we find that the WTR (Fin) and the HInvQ are able to predict risky behavior (see column 10). In the last column of Table 4, investigating the relationship between BMI and risk attitude, we find that higher risk-seeking behavior is correlated to higher BMI which is statistically significant for the HInvQ measure.

2.4.2 Predictive Ability of Single-item Risk Measures

While Section 2.4.1 focuses on risky behavior, Section 2.4.2 will focus on the outcomes per risk item. Considering the 11 x 7 matrix in Table 4, it becomes obvious that all items have some predictive ability - each of them has at least two significant coefficients. There are, however, larger differences. Some risk items have more strengths and others more weaknesses in explaining risky behavior. The

WTR (Gen), for instance, is only able to predict behavior in two out of our eleven cases, while the HInvQ is able to explain behavior in six cases. The two measures with more general ability are the HInvQ and the EG (No Loss) item. Overall, it can be said that the predictive ability of the single risk items is disappointing in our sample. In only 21 out of 77 efforts of predicting behavior, single-item risk measures are successful. The good news is, however, that despite limited predictive ability there is never a wrong prediction since all significant coefficients in Table 4 have the theoretically expected sign.

Of course the specific numbers in this exercise should not be over-interpreted because other kinds of risk behavior or another sample population may deliver different results (see Appendix section). Despite these unavoidable limitations, there seem to be three plausible and potentially robust lessons emerging: (1) It makes a difference which measure one takes (regarding explanatory power) due to heterogeneity across risk measures. (2) Our table indicates that there is an element of domain specificity in risk measurement. The WTR (Fin) explains "borrowing" well, while the HInvQ and GP are able to explain investment related risk behavior. However, domain specificity is not the only part of the overall structure we find because beyond these close relations, some risk measures are able to explain various kinds of risky behavior whereas other measures cannot. (3) An interesting side-aspect is that the more "precise" experimental measures do not show better predictive ability than the simple survey items (like the HInvQ) on average and in our setting.

Given these results, it seems plausible to combine several risk items into a latent risk item in order to reduce noise in single-item responses. In the next section, we construct an average across all risk tasks and test whether this multiple-item risk measure performs better than any single-item risk measure.

2.5 Predictive Ability of Multiple-Item Risk Measures

Section 2.5 reports findings on multiple-item risk measures. First, we describe the creation and performance of the 'all-item' risk measure (Section 2.5.1). Then we derive three factors from the seven risk measures (Section 2.5.2), and make use of this information by creating simplified measures based on three or two items (Section 2.5.3). In the next step, we show the efficiency gain in predicting behavior by building random two-item measures (Section 2.5.4). Finally, we search for principles in how to select risk-items for a powerful and parsimonious multiple-item risk measure (Section 2.5.5).

2.5.1 All-Item Risk Measure and Risky Behavior

In this section, we perform the same analysis as in Section 2.4, however, this time using the average of the seven measures to explain behavior. Averaging facilitates interpretation of results because it makes sure all variables contribute evenly to a scale when items are added together. We proceed in two steps: first, the mean of a specific risk item is subtracted from the value of each item, resulting in a mean of zero. Second, these different values are divided by the standard deviation – resulting in a standard deviation of one.

Regarding the 'all-item' risk measure, we only report core results (Table 4, bottom) while full tables can be provided upon request. We find that the all-item measure explains more risky behavior than any single-item measure. While the best items in Table 4 — the HInvQ and the EG (No Loss) task — are only able to explain risky behavior in six to five cases, our multiple-item risk measure is able to explain them in nine cases. We show this with regard to the eleven kinds of risky behavior.

Thus, the all-item measure appears to be more robust. However, for some specific forms of risky behavior our single-item measures predict better behavior. We assess the predictive ability by the level of significance reached. By this criterion, the average measure is best in predicting lottery expenditure as it reaches the 1%-significance level. This is better than the 5%-levels reached by WTR (Gen) (see Table 4, column 1). We can say here that a one-standard deviation increase in risk-tolerance (of the average measure) increases the share of lottery expenditure by approximately 4%. Regarding future lottery expenditure, the average measure is as good as EG (No Loss) because both risk measures realize coefficients at the 1%-level. Regarding all other nine kinds of risky behavior, however, the average measure is never best, although it is significant in seven of the remaining nine cases. This indicates – from another perspective – that there seems to be a trade-off between robustness and precision in specific cases.

For practicality reasons, it would be interesting to see whether we can reduce the number of items considered and still keep the power in explaining behavior. For this, we employ a factor analysis to reveal a reduced set of factors from the seven risk measures. We use these factors to predict risky behavior, learn about their explanatory power and find out if there might be substitutes for them.

2.5.2 Factor Analysis and Risky Behavior

In the correlation matrix, we saw varying degrees of correlation across risk elicitation methods. We aim now to reduce the number of variables and to detect a structure in the relationship between risk items, employing a standard factor analysis. We are able to extract three relevant factors, after careful examinations of eigenvalues, the proportion of variance explained, and scree plot criterion (Details can be found in Appendix E). The first factor accounts for 43% of the variance and the loading is dominated by the HInvQ. The second factor explains 32% of all variance and is dominated by the two EG items. The third factor is dominated by the WTR (Gen) and to some extent by WTR (Fin) and has the lowest explained variance (25%).

Next, we consider each factor and its power in explaining the same behavioral items as above (Table 5). We show statistically significant coefficients only, employ the same control variables and report clustered standard-errors. We find that Factor 1 has – as to be expected – the best explanatory power among the three factors as it explains risky behavior in 7 out of 11 cases. It is, for instance, able to explain investment and borrowing behavior at the 10% level, i.e. two behavioral items from the financial domain. We also find the expected signs for the domains of risk avoidance and behavior towards risk. Thus, we find slight domain-specificity.

Factor 2, which captures the two EG items, is able to explain risky behavior also in the domain of risk avoidance but also in other areas than Factor 1. We can see that it explains future gambling behavior well (at the 5% significance level). Factor 3, relying on the two WTR items, is only able to explain gambling behavior at the 5% significance level.

Using factor analysis, we find that we can reduce our seven risk items to three factors. Each of the three factors captures the items which are closely related to each other (i.e. the two WTR or the EG items). Moreover, running a regression with each factor as an explanatory variable with risky behavior, we find that the three factors are capturing different dimensions of risky behavior. However, it seems difficult — or at least highly speculative — to relate the factors to intuitive dimensions of risk attitude.

2.5.3 Reduced Multiple-Item Risk Measures and Risky Behavior

We saw in the previous section that the factor analysis yield three factors which load on a very limited set of risk attitude items. This motivates to extract one item, each one from the three factors, i.e. the HInvQ, the EG (No Loss) task and the WTR (Gen). We then average these three items by following the same methodological procedure as in Section 2.5.1.

By applying this reduced multiple-item risk measure on the 11 areas of risky behavior, we find that it is less robust in explaining behavior compared to the all-item risk measure. This may be reasonable since it contains less variance and information on risk attitudes compared to the seven measures (Table 5, bottom). Yet it is able to predict seven risky behaviors which is still more than using the best of the single-item risk measures. In other words, despite lacking the same strength as the average of seven items, using an average of three items has still more predictive ability than using any single measure alone in order to reveal risky behavior.

Since sometimes even three items can be infeasible in the field, we reduce our three-item risk measure further by including only two risk items; the HInvQ and the EG (No Loss) item. Table 5 (bottom) shows the predictive power of this reduced multiple-item risk measure with two items. We find that the multiple-item risk measure is as good as the HInvQ in explaining risky behavior and better than all other single-item measures by explaining six risky behavioral items out of eleven.

Yet, despite lacking the same strength as the average of the all-item measure, using an average of three or two items is still more fitting and robust in predicting risky behavior than using any single-item risk measure alone.

2.5.4 Efficiency Gain from Creating Random Two-Item Risk Measures

As a consequence from the last section, we now examine the average gain one can expect by moving from a single-item to a two-item risk measure. Thus, we assume the viewpoint of someone who does not know ex ante which measure may be best to explain a certain type of behavior but has a pool of seven measures from which he can select.

Regarding the single-item risk measures, we know that they can explain between 2 and 6 forms of risky behavior out of 11 cases. We are now interested to see the outcome if we combine without any prior restriction single-item measures together. Table 6 shows the result: WTR (Gen), for example, can explain 2 kinds of risky behavior, combining it with WTR (Fin) worsens the outcome since only 1 kind of risky behavior can be explained, combining it with HInvQ leads to 5 successes, combining it with GP leads to 3 successes etc. On average, one can expect an improvement from 2 kinds (for the WTR (Gen) single-item risk measure) to 3 kinds of risky behavior (for any combination of WTR (Gen) with a second single-item risk measure), although within a range of 1 to 5. Continuing this exercise for the other six single-item risk measures shows the average efficiency gain from taking two items instead of one: this move will improve the explanatory power from 3.0 to 4.1 cases, i.e. by more than 30%.

In a further analysis, we conduct the same procedure as above but this time combining any three risk items to a multiple-item risk measure. We find that our 'any three-item risk measure' is able to predict 5 risky behavior on average (Results available upon request). Compared to the 'any two-item' risk measure, the improvement is about another 25% (from 4.1 to 5).

The analyses above demonstrate the gain one may expect when applying any two or three-item risk measure. Gains can be, of course, much higher if one knows ex ante about the predictive power of specific measures and how these relate to each other.

2.5.5 Principles for Building Multiple-Item Risk Measures

It is our ambition to create multiple-item risk measures which have high explanatory power in predicting risky behavior. In the following, we consider three characteristics of our risk measures and analyze their impact on the performance of multiple-item risk measures.

Similar Framings. The first characteristic we are interested in, is the effect of framing, that is, when decision makers respond differently to distinct but objectively equivalent descriptions of the same problem. Most studies on framing effects describe logically equivalent decision situations in either a positive or a negative light. The "Asian Disease Problem", due to Tversky and Kahneman (1981), is a well-known example. Describing a choice between medical programmes in terms of lives saved or lives lost led to dramatically different answers, although the problems were logically equivalent. Levin et al. (1998) provide an overview and classification of framing effects found in numerous laboratory and questionnaire studies.

We analyze this issue here from another perspective. Can we improve the predictive ability of our risk measures if we combine several single-item measures which are similarly framed?

The seven single-item risk measures include two cases where measures are very similarly framed. First, we have the case of the two EG measures which use the same framing but differ regarding the payoff structure, in particular since one experiment considers potential losses. Second, there are the two survey items WTR (Gen) and WTR (Fin) which use the identical frame but refer to different domains. Thus, we keep items with similar framings but vary the domain.

We combine now for both cases the two single-item risk measures in the same way as described in Section 2.5.1. Using the new measures in predicting risk behavior, we see that these multiple-item risk measures perform disappointingly (Table 7, Panel A). The combined EG measures are able to predict 4 risky behaviors, compared to 2 for EG (Loss) and 4 for EG (No Loss), thus we can clearly infer that there is no real advantage. The combined WTR measures perform even worse than the single-items because the latter explain 2 behaviors each but the combination of the latter succeeds only once.

Overall, our data suggest that using multiple-item risk measures with the same framing does lead to a rather weak multiple-item risk measure.

Similar Domains. The second characteristic we try to investigate is whether averaging single-item measures in similar domains improves predictive ability. Employing similar domains but different framings, allows us to use four cases – all from the financial domain: (1) HInvQ and GP, (2) HInvQ and WTR (Fin), (3) GP and WTR (Fin), and (4) all three items together.

Results indicate that combining single-item risk measures from the same domain can be useful (see Table 7, Panel B): in case (1) the new measure explains 7 behavioral variables (6 and 3, respectively, as single-item risk measures), in case (2) the new measure explains 4 (instead of 6 and 2), in case (3) it explains 5 (3 and 2 before) and in case (4) it explains 8 behavioral items (6, 3 and 2 before).

We conclude that combining items within the same domain but using different framings may increase predictive power.

Repetition. The third characteristic we analyze, is the reliability of responses over time for the same measure since the WTR (Gen) had been asked a few months before the experimental survey in August 2013. We find that creating a multiple-item risk measure with repeated risk items is able to explain 1 kind of behavior (lottery expenditure), whereas the non-repeated WTR (Gen) could explain 2 behavioral items (see Table 7, Panel C). It becomes obvious that the reduction of noise by repeating the same risk experiments does not seem to be crucial.

Summarizing the analysis of this section, the most important principle for building a powerful multiple-item risk measure is using items with different framings within the same domain. We do not find evidence that using items from different domains would be generally helpful and also reducing noise by averaging over time does not contribute to an overall progress within our data.

2.6 Conclusion

For researchers who try to use a risk elicitation method as an explanatory variable, in household surveys for instance, it is important to know ex ante which risk item to choose from, particularly given the high cost involved in the data collection process. Unfortunately, the few available studies converge into the insight that risk elicitation methods do not capture exactly the same. They may be related to each other, however, to such a limited degree that the choice of any measure can lead to divergent conclusions. Our paper attempts to address this issue.

This is possibly the first study in which a comprehensive external validation of various risk elicitation methods has been conducted. We show in our study that risk measures are not only different but that they also differ in their explanatory power. They have specific abilities in explaining particular forms of risky behavior. Some measures seem to perform generally better than others, at least with respect to our sample population and overall design. We find that a large part of observed heterogeneity seems to come from noise in measurement and one way to reduce this is by taking the average across all seven risk elicitation methods.

Our resulting multiple-item risk measures — using the average of two, three or all seven risk items – offer a behaviorally meaningful alternative in revealing preferences in a reliable way. Our finding shows that such multiple-item risk measures perform better than any single-item risk measure. In many detailed analyses we support this finding and show that it is robust to several concerns (for more details on robustness tests see Appendix F). Our results also suggest that employing risk tasks with different framings increases predictive ability of the multiple-item risk measure. Therefore, our study not only informs us about the behavioral validity of each item but also offers an alternative for a reliable and behaviorally valid risk measure. This also contributes, for instance, to the improvement of practical techniques aimed at the accurate elicitation of risk attitudes in surveys.

Nevertheless, this is just a single study and it would be interesting to learn from additional research whether this main finding largely holds and in which ways it could be improved. For all kinds of applications the improvement on simple and at the same time reliable measures of risk attitude is urgent.

Table 1: Descriptive Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Panel A: Individual Characteristics					
Female	0.575	0.495	0	1	764
Age	54.496	12.465	17	79	764
Height	158.148	7.641	140	185	749
Years of Schooling	5.674	3.113	1	17	737
Marital Status	0.831	0.375	0	1	751
Household Size	4.102	1.722	1	12	749
Log Per Capita Consumption	7.569	0.599	5.98	10.331	749
Panel B: Behavioral Variables					
Basic Algebra	3.632	1.311	0	6	763
Farmer	0.655	0.476	0	1	760
Self-Employed	0.083	0.272	0	1	761
Lottery Spending (Past)	630.327	2407.258	0	36000	759
Lottery Spending (Future)	28.880	172.350	0	2934	759
Amount of Investment	5083.793	12215.916	0	124496	764
Plan to Invest	0.473	0.500	0	1	764
Borrowing (General)	0.713	0.452	0	1	764
Borrowing (Business)	0.082	0.275	0	1	764
Risk-Mitigating Activities	0.496	0.500	0	1	764
Number of Risk-Mitigating Activities	0.928	1.213	0	8	764
Number of Insurance (General)	1.959	1.753	0	1	764
Number of Insurance (Health)	0.058	0.233	0	1	764
Number of Insurance (Accident)	0.076	0.265	0	1	764
Body Mass Index	23.067	3.727	12.889	39.184	742
Overweight	0.299	0.458	0	1	764

Notes:
Height is in cm. Household Size is the headcount of persons living in the household for at least 180 days. Log Consumption refers to the natural logarithm of consumption divided by OECD adult equivalents AE (AE = 1+0.7*(adults-1) + 0.5*children). Lottery Expenditure is the total annual lottery expenditure in the last 12 months. Future Lottery Expenditure is the expected lottery expenditure in the next drawing. Investment is amount of investment reported by the household. Results have been calculated in purchasing power parity adjusted US dollars. In February 2008 the International Comparison Program published purchasing power parities stating that 15.93 THB equal 1 PPP USD. Planned Investment is if the household plans any investment in the next five years. Risk-Mitigating measures indicate whether respondents took up any measures in order to prevent any future shocks/risks. Health and Accident Insurances are the voluntary health/accident insurance take-up. BMI is computed weight/height^2. Overweight are those having a BMI $>$ 25. We employ the subsample between 17-79 years old.

Table 2: Descriptive Statistics of Risk Eliciation Methods

Variable	Mean	Std. Dev.	Min.	Max.	N
General Willingness to take risk (WTR Gen)	6.85	3.02	0	10	764
Financial Willingness to take risk (WTR Fin)	6.47	3.28	0	10	764
Hypothetical Investment Question (HInvQ)	50.88	21.38	0	100	764
Certainty equivalent (CEquiv)	7.94	7.14	1	20	763
Eckel-Grossman with Loss (EG Loss)	3.18	1.56	1	5	763
Eckel-Grossman without Loss (EG No Loss)	3.03	1.49	1	5	764
Gneezy-Potters (GP)	36.37	30.56	0	100	764

Notes:
This table shows the estimates of all risk elicitation methods. The general willingness to take risk item asks on an 11-point Likert scale "Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?" The financial willingness to take risk, also an 11-point Likert scale, asks "When thinking about investing and borrowing are you a person who is fully prepared to take risk or do you try and avoid taking risk?" The hypothetical question asks "Imagine you just won 100 000 Baht in a lottery and you can invest this money in a business. There is a 50% chance that the business is successful. If the business is successful you double the amount invested after one year. If it is not successful you will lose half the amount you invested. What fraction of the 100 000 Baht would you invest in the business?" Results are divided by 1000. The certainty equivalent task is an experiment with a lottery that offers a lottery with a 50-50% chance of receiving either 0 or 300 THB and alternatively a safe payoff of 0 THB. Each row, the value of the safe payoff is increased by 10 THB. In the Eckel-Grossman experiment subjects must play out one of five possible gambles. All the gambles involve a 50/50 chance. The EG with loss treatment involves a negative payoff of -30 THB. In the Gneezy-Potters experiment subjects have to decide to allocate a fraction of 100 THB in a risky business or to keep it.

Table 3: Spearman's rank correlations across elicitation methods

	WTR (Gen)	WTR (Fin)	HInvQ	CEquiv	EG (Loss)	EG (No Loss)	GP
WTR (Gen)	1.000						
WTR (Fin)	0.359***	1.000					
	(0.000)						
HInvQ	0.086**	0.122***	1.000				
	(0.018)	(0.001)					
CEquiv	0.034	0.000	0.083**	1.000			
	(0.356)	(0.998)	(0.022)				
EG (Loss)	0.094**	0.027	0.063*	0.100***	1.000		
	(0.010)	(0.451)	(0.082)	(0.006)			
EG (No Loss)	0.031	-0.014	0.008	0.074**	0.436***	1.000	
	(0.398)	(0.695)	(0.820)	(0.042)	(0.000)		
GP	0.030	0.046	0.201***	0.030	0.078**	0.098***	1.000
	(0.404)	(0.203)	(0.000)	(0.405)	(0.032)	(0.007)	
N: 760							

Notes:
The table reports pairwise Spearman rank correlation coefficients for the subsample with age of 17-79. Statistical significance is in parenthesis. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table 4: Single and Multiple-Item Risk Measures

	(1) Lottery (Expend/Cons)	(2) Future Lottery Expenditure	(3) Self-employment	(4) Farming	(5) Investment (Expend/Income)	(6) Plan to Invest	(7) Borrowing	(8) Risk Mitigating	(9) Number of Insurance	(10) Health Insurance	(11) BMI
WTR (Gen)	0.011**		0.007**								
	(0.00)		(0.00)								
WTR (Fin)							0.021***			-0.005**	
							(0.01)			(0.00)	
HInvQ			0.001**		0.008*		0.002*	-0.002***		-0.001***	0.014**
			(0.00)		(0.00)		(0.00)	(0.00)		(0.00)	(0.01)
CEquiv				-0.009***		0.008***					
				(0.00)		(0.00)					
EG (Loss)	0.027***	10.570**									
	(0.01)	(4.46)									
EG (No Loss)		16.073***		-0.023*				-0.027*	-0.103**		
		(5.73)		(0.01)				(0.01)	(0.04)		
GP		0.519**	0.001**		0.007***						
		(0.02)	(0.01)		(0.00)						
Observations	711	710	715	710	715	715	708	715	715	715	710
Average of 7 Items	0.042***	12.893***	0.019*	-0.035*	0.120*	0.046**	0.040**	-0.053**		-0.019**	
	(0.02)	(4.53)	(0.01)	(0.02)	(0.07)	(0.02)	(0.02)	(0.02)		(0.01)	
Observations	709	709	713	709	713	713	713	713		713	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Estimator	OLS	OLS	Probit	Probit	OLS	Probit	Probit	Probit	OLS	Probit	OLS

Notes:
The dependent variables are the behavioral variables from the household survey. Expend/Cons in parenthesis is the total amount of household expenses in the last 12 months as a share of total consumption. Expend/Income in parenthesis is the total amount of household expenses in the last 12 months as a share of total income. See Appendix B for further details. Controls include Female, Age, Height, Marital Status, Household Size, Education and Log Per Capita Consumption. We employ the subsample with age of 17-79. Clustered errors on the village level are in parenthesis. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table 5: Factor Analysis and Multiple-Item Risk Measures

	(1) Lottery (Expend/Cons)	(2) Future Lottery Expenditure	(3) Self-employment	(4) Farming	(5) Investment (Expend/Income)	(6) Plan to Invest	(7) Borrowing	(8) Risk Mitigating	(9) Number of Insurance	(10) Health Insurance	(11) BMI
Factor 1			0.020**		0.176*		0.041*	-0.056***	-0.110*	-0.030***	0.308**
			(0.01)		(0.09)		(0.02)	(0.02)	(0.07)	(0.01)	(0.14)
Factor 2		28.750**						-0.047*	-0.169*		
		(10.99)						(0.03)	(0.09)		
Factor 3	0.033**										
	(0.02)										
Observations	709	709	713		713		713	713	713	713	709
Average of 3 Items	0.046**	17.07***	0.137*		0.120*			-0.167***		-0.144*	0.318**
	(0.02)	(5.47)	(0.08)		(0.07)			(0.06)		(0.08)	(0.14)
Average of 2 Items		27.33**			0.153*			-0.161***	-0.108**	-0.161**	0.285**
		(11.34)			(0.08)			(0.05)	(0.05)	(0.08)	(0.13)
Observations	711	711	715		715	108		715	715	715	710
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Estimator	OLS	OLS	Probit	Probit	OLS	Probit	Probit	Probit	OLS	Probit	OLS

Notes:
The dependent variables are the behavioral variables from the household survey. Expend/Cons in parenthesis is the total amount of household expenses in the last 12 months as a share of total consumption. Expend/Income in parenthesis is the total amount of household expenses in the last 12 months as a share of total income. See Appendix B for further details. Controls include Female, Age, Height, Marital Status, Household Size, Education and Log Per Capita Consumption. We employ the subsample with age of 17-79. Clustered errors on the village level are in parenthesis. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table 6: Average of Any Two Single-Item Risk Measures

	WTR (Fin)	HInvQ	CEquiv	EG (Loss)	EG (No Loss)	GP	Average of Any Two Items	Average of Single-Item
WTR (Gen)	1	5	4	2	3	3	3.0	2
WTR (Fin)		4	5	3	3	1	2.80	2
HInvQ			7	6	9	6	6.16	6
CEquiv				3	4	3	4.33	2
EG (Loss)					4	4	3.66	2
EG (No Loss)						4	4.50	4
GP							3.50	3
Average							4.13	3
N: 760								

Notes:
The table reports significant results for any two multiple-item risk measure in explaining risky behavior. We follow the same procedure as in Table 4. We employ the subsample with age of 17-79. Average denotes the average power of any two-multiple risk item in explaining a number of risky behavior.

Table 7: Single-Item Risk Measures and Risky Behavior (Framing, Domain, Repetition)

	(1) Lottery (Expend/Cons)	(2) Future Lottery Expenditure	(3) Self-employment	(4) Farming	(5) Investment (Expend/Income)	(6) Plan to Invest	(7) Borrowing	(8) Risk Mitigating	(9) Number of Insurance	(10) Health Insurance	(11) BMI
					A: Framing						
EG (Loss)*EG (No Loss)	0.029*	24.16***						-0.033*	0.134*		
	(0.02)	(0.07)						(0.02)	(0.07)		
WTR (Fin)*WTR(Gen)								0.039**			
								(0.02)			
					B: Domain						
HInvQ*GP	0.047**	21.12**	0.027***		0.250***	0.0102*		-0.050***		-0.026***	
	(0.02)	(9.21)	(0.00)		(0.09)	(0.06)		(0.02)		(0.01)	
WTR (Fin)*GP			0.021**		0.139**	0.038*	0.049***			-0.019***	
			(0.01)		(0.06)	(0.02)	(0.02)			(0.00)	
WTR (Fin)*HInvQ						0.041**	0.077***	-0.039**		-0.029***	
						(0.02)	(0.02)	(0.02)		(0.01)	
WTR (Fin)*HInvQ*GP	0.039**	8.746*	0.025***		0.194**	0.048**	0.058***	-0.044**		-0.030***	
	(0.02)	(5.10)	(0.01)		(0.09)	(0.02)	(0.02)	(0.02)		(0.01)	
					C: Repetition						
WTR (Apr)*WTR (Aug)	0.065***										
	(0.02)										
HInvQ (Apr)*HInvQ(Aug)	0.054*					0.046*	0.045**			-0.026**	
	(0.03)					(0.03)	(0.02)			(0.01)	
Observations	709	709	713	709	713	713	713	713		713	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Estimator	OLS	OLS	Probit	Probit	OLS	Probit	Probit	Probit	OLS	Probit	OLS

Notes:
The dependent variables are the behavioral variables from the household survey. Expend/Cons in parenthesis is the total amount of household expenses in the last 12 months as a share of total consumption. Expend/Income in parenthesis is the total amount of household expenses in the last 12 months as a share of total income. See Appendix B for further details. Controls include Female, Age, Height, Marital Status, Household Size, Education and Log Per Capita Consumption. Apr denotes the item from the spring survey while Aug is the survey conducted in the summer. We employ the subsample with age of 17-79. Clustered errors on the village level are in parenthesis. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Appendix A Information about Risk Elicitation Methods

A.1 Implementation of Risk Eliciation Tasks

Implementation of the risk elicitation methods contained two village visits per day, one in the morning and one in the afternoon. We cannot fully eliminate the possibility that information had spread between villages; yet this is rather unlikely because most of the villages are far away from each other (18km on average).

The study was carried out by local enumerators with one of the research fellows being present at all times to ensure compliance. Some enumerators were different from those conducting the household survey but had extensive interviewer skills acquired in other surveys. We do not find any systematic interviewer fixed effects. The survey was translated from English into Thai and vice versa and was cross-checked by a Thai economics professor to avoid semantic difficulties. The interviewer training lasted for a total of five days. During these five days, a pilot study was conducted in three villages. We interviewed 830 individuals in total.

In general, enumerators were instructed to select the household member (usually the household head) who was previously interviewed in the household survey to participate in the experimental study. In case that person was not available, enumerators selected the closest family member present. In 44 cases we interviewed households which had not participated at the household survey, so we miss baseline information about these households (and its members) and drop them from our current analysis. Further, we restrict our sample to respondents aged between 17 and 79 years. It is assumed that respondents with age above 80 or below 17 may have more difficulties in understanding the experiments. Hence, we drop another 26 observations. Ultimately, we work with 760 observations.

The experimental sessions were conducted in the village town hall. To avoid observation, we made sure that respondents were separated across the town hall. Furthermore, decision spillovers are unlikely because individuals responded at different pace levels. Of course, we cannot exclude the possibility of observation altogether.

Upon arrival, the experimenter reminded participants of the confidentiality of the data. In order to ensure incentive compatibility, subjects are informed that after the experiment a random device will determine which experiment will be paid out depending on their decision. The maximum number of participants in any session was 10. Care was taken to ensure that subjects understood the decisions they were to make.

Once all seven choices were made, one decision was randomly chosen from the incentivized part for payment. The respondents had to pick a number from a non-transparent bag to determine which experiment is played out and a coin was used to determine the outcome of the risk game. Average earnings were 150 Thai Baht (THB), i.e. approximately 4 Euros slightly less than a one-day salary of an unskilled worker. The show-up fee was 50 THB (approx. 1 Euro). The EG (Loss) included a negative outcome (-30 THB). We, however, avoided negative payoff by providing an initial fee of 30 THB equal to the maximum loss that could be incurred due to ethical reasons in a manner similar to the Hey and Orme (1994) replication exercise done in Harrison and Rutström (2008, p.164). This is not given in the no loss experiment. It should be noted that this has the potential drawback that loss aversion might be underestimated. However, there is little evidence that the house money effect is likely to change the result when we compare the baseline study of Eckel and Grossman (2008) and our results. Each session included exactly the same set of instructions and was implemented in the same order. While the risk experiments took half an hour, the entire risk survey from the beginning to the final payoff took approximately two hours to complete.

A.2 Sampling Procedure of the Household Survey

Our risk survey is administered as part of a larger household survey which collects data from approximately 2,200 households in three provinces in Thailand. The household selection process follows a three-stage stratified sampling procedure where provinces constitute strata and the primary sampling units are sub-districts. Within each province, we exclude the urban area around the provincial capital city and confine the sample to the remaining rural areas. Within each sub-district, two villages are chosen at random. In the third stage, a systematic random sample of ten households was drawn from household lists of the rural census ordered by household size. Overall, the sampled households are representative for the rural areas in the considered provinces. Compared to the household survey which ran in three provinces, our risk survey was conducted in the province of Ubon Ratchathani only, the largest of the three provinces in Northeastern Thailand.

Given the sampling procedure, we control for within-cluster error correlation. Instead of randomly drawing individuals from the entire population, costs are reduced by sampling only a randomly-selected subset of primary sampling units (sub-districts), followed by a stratified selection of people within the chosen sampling units (village level). Hence, we cannot assume that observations within villages have uncorrelated errors. Households of the same village might be more similar on a wide variety of measures than are households that are not part of the village. We, therefore, cluster the standard errors at the village level (Cameron and Miller, 2013).

A.3 Description of Risk Elicitation Methods

In our data collection process for the experiment, we tried to keep the enumerator instructions as short and simple as possible in order to facilitate the understanding.

1. Self-reported risk attitude: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risk? (Please choose a number on

a scale from 0 to 10).

2. Attitudes towards risk change in different situations. When thinking about investing and borrowing are you a person who is fully prepared to take risk or do you try and avoid taking risk? (Please choose a number on a scale from 0 to 10).

3. Imagine you just won 100 000 Baht in a lottery and you can invest this money in a business. There is a 50% chance that the business is successful. If the business is successful you double the amount invested after one year. If it is not successful you will lose half the amount you invested. What fraction of the 100 000 Baht would you invest in the business?

4. Certainty equivalent experiment: This is game 1. It has 20 rows. In each row a decision has to be made. In each row we would like you to choose option A or option B. Option A is a certain amount of THB. It starts with 0 and goes up by 10 THB in every row. Option B is a lottery where a coin is thrown. If 'King' falls you win 300 Baht. If 'Palace' falls you get nothing. *(Enumerator shows the coin).* Please make your choice of Option A or B for each row. If this game is selected to be played with real money, you will be asked to draw a number from a bag. The bag contains the numbers 1 to 20 for the 20 rows. We will play with real money according to your choice. For example: If you draw the number X (Enumerator ID) from the bag, we play the game at this row for money. That means: If you chose option A you will receive (THB). If you chose option B we will toss a coin. If 'King' you win 300 Baht. If 'Palace' you win nothing.

Table A1: Certainty Equivalent Task

Row	**Option A**	**Tick Box**	**Tick Box**	**Option B**
1	0			300 : 0
2	10			300 : 0
3	20			300 : 0
4	30			300 : 0
5	40			300 : 0
6	50			300 : 0
7	60			300 : 0
8	70			300 : 0
9	80			300 : 0
10	90			300 : 0
11	100			300 : 0
12	110			300 : 0
13	120			300 : 0
14	130			300 : 0
15	140			300 : 0
16	150			300 : 0
17	160			300 : 0
18	170			300 : 0
19	180			300 : 0
20	190			300 : 0

5. This is game 2. There are 5 options. Please choose the one option that you would like to play the most. In each of the five options we flip a coin to determine the real money payoff. *(Enumerator shows coin)*. Please see the table on the showcard:
   In option 1 you win 50 Baht if King falls and 50 Baht if Palace falls.
   In option 2 you win 90 Baht if King falls and 30 Baht if Palace falls.
   In option 3 you win 130 Baht if King falls and 10 Baht if Palace falls.
   In option 4 you win 170 Baht if King falls and -10 Baht if Palace falls.
   In option 5 you win 210 Baht if King falls and -30 Baht if Palace falls.
   Now we ask you to make your decision. Which of these 5 options do you prefer? *(Enumerator: Please tick box!)*

Table A2: Eckel-Grossman with Loss Treatment

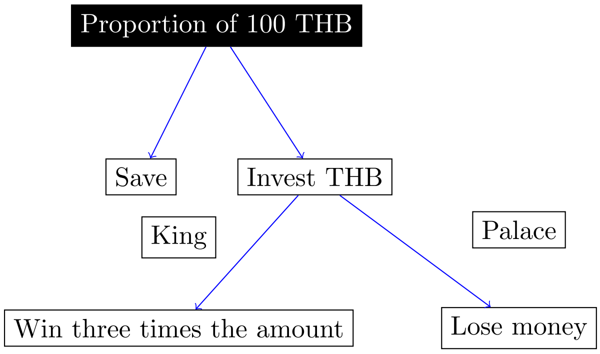
	Choice	Coin	Stakes	Tick Option
Option	1	King Palace	50 50	
Option	2	King Palace	90 30	
Option	3	King Palace	130 10	
Option	4	King Palace	170 -10	
Option	5	King Palace	210 -30	

6. This is game 3. There are 5 options. Please choose the one option that you would like to play the most. In each of the five options we flip a coin to determine the real money payoff. *(Enumerator shows coin)*. Please see the table on the showcard:
   In option 1 you win 80 Baht if King falls and 80 Baht if Palace falls.
   In option 2 you win 120 Baht if King falls and 60 Baht if Palace falls.
   In option 3 you win 160 Baht if King falls and 40 Baht if Palace falls.
   In option 4 you win 200 Baht if King falls and 20 Baht if Palace falls.
   In option 5 you win 240 Baht if King falls and get nothing if Palace falls.
   Now we ask you to make your decision *(Enumerator: Please tick box!)*. Which of these 5 options do you prefer?

Table A3: Eckel-Grossman without Loss Treatment

	Choice	Coin	Stakes	Tick Option
Option	1	King	80	
		Palace	80	
Option	2	King	120	
		Palace	60	
Option	3	King	160	
		Palace	40	
Option	4	King	200	
		Palace	20	
Option	5	King	240	
		Palace	0	

7. This is game 4. We offer you 100 Baht. There are two options for this money: you can keep money for certain or you can use money to play a game. We ask you to decide how much of the 100 Baht you want to use for these two options each. You can split the money in any way between these two options.



Appendix B Description of Variables

This section displays details of the variables we use in our regression.

B.1 Individual Characteristics

Female is a dummy variable. It takes the value 1 for female and 0 for male.

Age is the respondents' age in years.

Height is respondents' height in cm.

Years of Schooling is denoted as years in eduation.

Marital Status is a dummy variable. It takes the value 1 if married and 0 otherwise.

Household Size is the headcount of persons who are living in the household for at least 180 days.

LPCC is the log per capita consumption and it refers to the natural logarithm of household consumption per day divided by OECD adult equivalents AE (AE = 1+0.7*(adults-1) + 0.5*children).

B.2 Behavioral Variables

Farmer is a dummy variable. It takes the value 1 for being a farmer and 0 otherwise.

Self-employed is a dummy variable. It takes the value 1 for being self-employed and 0 otherwise.

Lottery Expenditure is counted by the total amount of household expenses for lotteries in the last 12 months.

Future Lottery Expenditure is denoted the question: How much money will you spend for the next drawing of the state lottery?

Investment is counted by the total amount of household purchases above 5000

THB and used for longer than a year since 2011.

Planned Investment constitutes the question: Do you plan to invest in the next five years in the agricultural/non-agricultural business?

Borrowing is a dummy variable. It takes the value of 1 if the household borrowed since 2011 and 0 otherwise.

Borrowing for Business is a dummy variable. It takes the value 1 if respondent applied for a credit for business purposes and 0 otherwise.

Risk-Mitigating Measures is a dummy variable. It takes the value of 1 if the household implemented risk measures since 2011 and 0 otherwise.

Number of Risk Measures is number of actual implemented risk-mitigating measures since 2011.

Number of Insurances is the total sum of voluntary insurances of the household.

Health Insurance is a dummy variable. It takes the value of 1 if the household took up additional health insurance next to the public one and 0 otherwise.

Accident Insurance is a dummy variable. It takes the value of 1 if the household took up an additional accident insurance and 0 otherwise.

Body Mass Index is computed by weight/height^2.

Overweight is a dummy variable. It takes the value of 1 if the person has a BMI > 25. The WHO definition for BMI is if a person has a BMI greater than or equal to 25.

B.3 Cognitive Aptitude

We asked five basic algebra questions and tested additionally their word fluency.

1 What is 45 + 72?

2 You have 4 friends and you want to give each friend 4 sweets. How many sweets do you need?

3 What is 5% of 200?

4 You want to buy a bag of rice that costs 270 Baht, You only have one 1000 Baht note How much change will you get?

5 In a sale, a shop is selling all items at half price. Before the sale a mattress costs 3000 Baht. How much will the mattress cost in the sale?

6 A second-hand motorbike dealer is selling a motorbike for 12000 Baht. His is two thirds of what it costs new. How much did the motorbike cost new?

Word fluency: I would like you to name as many different animals as you can in 60 seconds *(Enumerator counts with a stopwatch)*.

Appendix C Risk Measures and Socio-economic Correlates

Table C.1 explores the relationship between various risk elicitation methods and socio-economic correlates. In explaining the individual risk attitude, we rely on a set of seven standard variables which are potential determinants of risk attitude (Dohmen et al., 2011). Noticeable is the reduction of the sample size from 760 to 715 which is due to missing observations in terms of schooling and consumption. We shortly comment on significant coefficients only.

We find that women tend to make more risk-averse choices, although this turns significant in two out of seven measures only. This finding is consistent with an emerging literature documenting differences in the risk preferences of men and women using various risk elicitation tasks (Croson and Gneezy, 2009). Age has an ambiguous relationship with risk attitude. It seems that younger people are more risk seeking in the WTR (Fin) as well as the HInvQ. This is in line with the results found in (Dohmen et al., 2011). This result is different in both EG tasks where older people make more risk-seeking choices. For marital status we find in three cases statistically significant evidence that married people make more risky choices. More schooling seems to increase risk tolerance. With regard to remaining variables, household size and higher consumption tend to go along with more risk tolerance which is statistically significant in one case only. Overall, significant relations have the expected signs; with the exceptions of age and height.

Table C.1: Determinants of Risk Measures

	(1) WTR (Gen)	(2) WTR (Fin)	(3) HInvQ	(4) CEquiv	(5) EG (Loss)	(6) EG (No Loss)	(7) GP
Female	0.395	-0.005	-4.183**	-1.225*	-0.054	0.000	-4.556
	(0.28)	(0.28)	(1.98)	(0.66)	(0.17)	(0.15)	(2.78)
Age	0.011	-0.024*	-344.2***	-0.015	0.013*	0.017***	-0.196
	(0.01)	(0.01)	(0.07)	(0.02)	(0.01)	(0.01)	(0.12)
Height	0.010	-0.006	0.169	-0.059	-0.002	0.007	-0.304*
	(0.02)	(0.02)	(0.12)	(0.04)	(0.01)	(0.01)	(0.17)
Years of Schooling	0.057	0.002	0.069	0.111	0.033*	-0.025	0.421
	(0.04)	(0.05)	(0.28)	(0.11)	(0.02)	(0.02)	(0.40)
Marital Status	0.772**	0.855**	6.912***	0.583	0.070	-0.023	4.238
	(0.35)	(0.37)	(2.27)	(0.76)	(0.15)	(0.14)	(3.34)
Household Size	0.004	0.126*	0.331	0.133	-0.050	-0.011	0.609
	(0.08)	(0.07)	(0.44)	(0.19)	(0.04)	(0.04)	(0.74)
LPCC	0.015	0.205	2.743**	0.320	-0.115	-0.076	3.260
	(0.25)	(0.20)	(1.36)	(0.46)	(0.10)	(0.10)	(1.96)
Constant	3.436	5.947	17.345	14.784*	3.743*	1.824	64.833**
	(3.65)	(3.72)	(21.64)	(7.85)	(2.06)	(1.61)	(31.66)
Observations	719	719	719	718	718	719	719
R-Squared	0.01	0.03	0.09	0.01	0.01	0.03	0.02
Estimator	OLS	OLS	OLS	OLS	OLS	OLS	OLS

Notes:
The dependent variables are risk elicitation methods. Columns [1-7] report estimates of least square estimations. Female is a dummy (1=yes, 0=no). Age is age of respondents' in years. Height is in cm reported by the respondent. Marital Status is a dummy (1=yes, 0=no). LGPP is the log per capita consumption. It refers to the natural logarithm of household consumption per day divided by OECD adult equivalents AE (AE = 1+0.7*(adults-1) + 0.5*children). We employ the subsample of 17-79 years. Clustered errors on the village level are in parenthesis. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Appendix D Theoretical Predictions of Risk Attitude and Risky Behavior

Gambling (Playing Lottery). We explore the relationship of risk attitude and lottery participation in our sample since expenditures for lottery and potentially other forms of gambling seems to be an important part of many poor households budgets. Poor households do not spend all of their money on food and non-food items but instead entertainment related activities, i.e., alcohol, tobacco, festivals and also playing lotteries account for an important share of many poor household's budgets (Banerjee and Duflo, 2007). The relationship between gambling and risk attitude is close because participation in a lottery is a risky decision. Hence, the purchase of lottery tickets is a good indicator of risk-seeking behavior. The relationship is studied in numerous works (Clotfelter and Cook, 1990).

Risky Employment. Entrepreneurship is another prominent example of risky behavior. Knight (1921) considered the factors which influence entry into self-employment and stressed the special role of the entrepreneur as operating under uncertainty and bearing risk of business failure. Since running a business is equivalent to the choice of a risky prospect, the less risk-averse will become entrepreneurs while the relatively risk-averse will prefer to be employees and work for a fixed wage.

In line with this theoretical prediction, the main body of empirical evidence on risk, uncertainty and entrepreneurship shows risk tolerance of entrepreneurs. For example, risk tolerance explains undiversified portfolio holdings (Moskowitz and Vissing-Jörgensen, 2002). However, in a recent study, Holm et al. (2013) find that entrepreneurs in China do not generally differ from other people when it comes to behavior under uncertainty.

Eliciting risk attitudes on a sample of farmers is worth being undertaken for several reasons. First, farmers are used in their professional life to take decisions under uncertainty (they face production, market, and environmental risks). Second, being a farmer in a poorer rural area is often the last resort of those who are not entrepreneurial or mobile. Third, there is currently no consensus in the agricultural economics literature on the level of farmer's risk aversion (Reynaud and Couture, 2012). We are also aware that the decision for being self-employed (farmer) and lower (higher) risk aversion suffers from reverse causality. We, therefore, interpret results as correlates.

Financial Behavior. First, standard portfolio theory predicts that the share of wealth an individual is willing to invest in risky assets depends on his/her degree of risk aversion. Even though exact conditions are hardly met in reality, in particular as poorer farmers do not and cannot hold a market portfolio, the more risk-averse should hold safer portfolios and make less risky investments.

Second, planning an investment in the future is embedded in uncertainty about the conditions under which the planned investment may take place. We hypothesize that risk-tolerant rather than risk-averse individuals should be more prone in planning to conduct considerable investments.

Third, borrowing can be generally seen as a decision which entails risk because the borrower has agreed to future repayment without knowing his future economic situation. Thus, the less risk-averse individuals should be more likely to borrow more (i.e. for investment purposes).

Risk Avoidance. Income risk is a central feature of rural areas in developing countries. A major concern is how well households are able to mitigate adverse effects of income risk on their welfare. In the absence of formal insurance markets, household undertake actions to reduce the variability of income. Risk-averse individuals may choose to implement or undertake more risk-coping mechanisms (i.e. substitute crops, diversify agricultural portfolio etc.) than a risk-seeking individual. The classical model of the demand for insurance elaborated by Mossin (1968) implies that risk-averse individuals should fully insure if insurance is offered at fair terms. If insurance is unfair, the amount purchased will depend on one's degree of risk aversion: the more risk-averse will demand more insurance coverage.

Behavior Towards Health. Regarding health issues we consider the case where next to the free health insurance from the state, respondents also chose to have an additional health insurance with better coverage. The expectation is about the same as for the number of insurance contracts analyzed before. In the field of health economics, attitudes toward risk are also likely to affect the propensity to engage in behaviors that either increase or decrease mortality risk, such as cigarette, smoking, or seat belt use. Viscusi and Hersch (2001) and Anderson and Mellor (2008), for example, use smoking status and seat belt use to control for risk preference in terms of employment-related risk-taking. We compute the BMI of our respondents and use this measure to relate risk attitude with health decisions. The measure of BMI is also used in the study of Sutter et al. (2013) who find that higher BMI of students is strongly associated with less risk aversion.

Appendix E Factor Analysis

In the correlation matrix we saw varying degrees of correlation across risk elicitation methods. We aim now to reduce the number of variables and to detect structure in the relationships between risk items, employing a standard factor analysis. The Kaiser-Meyer-Olkin (KMO) measure verifies the sampling adequacy for the analysis. All KMO values for individual items are above 0.5, supporting retention for the analysis. Factor analysis yields three factors with an eigenvalue of 1.16, 0.87 and 0.70. The first factor accounts for 43% of the variance and the loading are dominated by the HInvQ. The second factor explains 32% of all variance and is dominated by the two EG items. We can see that 75% of the common variance shared by the seven variables can be accounted by the first two factors. The third factor is dominated by the WTR (Gen) and to some extent by WTR (Fin) and has the lowest explained variance.

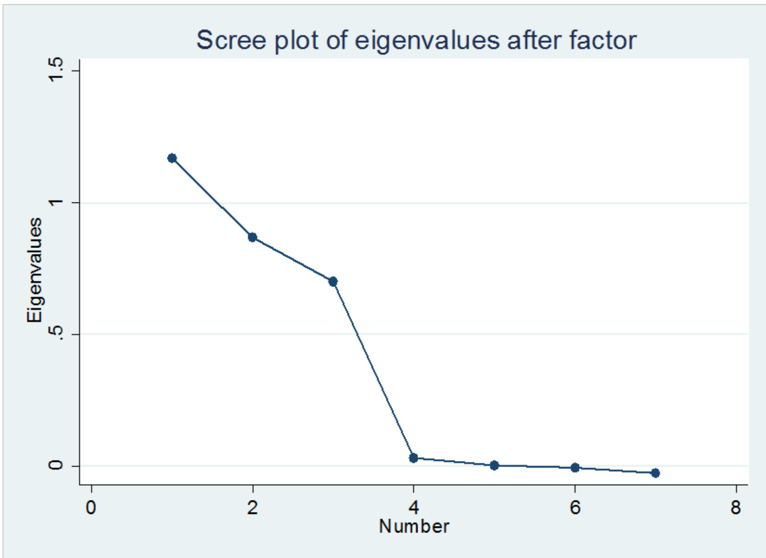
After careful examinations of eigenvalues, proportion of variance explained and scree plot criterion, three factors are identified for further use. After rotation, we find a clearer picture of the relevance of each variable in the factor. We drop those variables with a loading smaller than 0.30 (see Table E.1). Based on the a priori classifications, a clear and interpretable underlying structure is identified. We are able to confirm the result described above, in that Factor 1 is mostly defined by the HInvQ, Factor 2 is defined by the two EG experiments, while Factor 3 is defined by the two WTR items.

Table E.1: Factor Analysis

Variable	Factor 1	Factor 2	Factor 3
WTR (Gen)			0.7548
WTR (Fin)			0.4852
HInvQ	0.9591		
CEquiv			
EG (Loss)		0.6216	
EG (No Loss)		0.7085	
GP			

Notes:
Factor analysis pattern matrix. Rotation method is promax with Kaiser normalization. Table shows for each risk elicitation method the factor loadings that are greater than 0.3.

Figure E.1: Screeplot



Appendix F Robustness Tests

Robustness examinations cover four issues: We use further extensions of the behavioral items, including variations of consumption and income definitions (Section F.1). Motivation and results of three other robustness exercises are described in Sections F.2., F.3, and F.4.

F.1 Extension of Risky Behavior

(1) Since any set of risky behavior is arbitrary, we document here either slight variations of already used measures or alternatives which are available from the household questionnaire. Overall, we present results for another eight risky behavioral items.

We use the share of past lottery expenditure as part of income instead of consumption. (2) We use the log of investment in our robustness section (instead of investment quota). This seems reasonable because we have a large share of respondents who made large amount of investments while others made none. Hence, in order to reduce the variation caused by extreme values, we take the log of investment and test whether the results are robust. It must be noted, however, that we will lose some observations, as those who made no investments will drop out. In previous tables we used the dependent variable the share of household investment as part of their income. In Column (3) we use the total amount of investment while in column (4) we use the share of household investment as part of their consumption. We use variations of the investment variable to show that the relation between risk and investment is not random. (5) In addition to borrowing, we use borrowing for business. The household survey not only asks whether they borrowed any loans in the last two years but also for which purpose, whether it was to repay back existing loans, to pay for medical bills, school fees, but also for business purposes among other things. (6) The number of risk-mitigating measures instead of the indicator variable. (7) We also consider accident insurance (as opposed to health insurance). (8) As another proxy for BMI, we will only consider those who are overweight according to the WHO definition.

Table F.1 shows results which reproduce the former Table 4 with the exception that we take the alternative variables. At first glance, we see that in 17 out of 42 cases a single-item risk measure is able to explain behavior. Hence, the main picture drawn from Table 4 is not changed. However, looking more carefully, we see that there are slight changes in significant results. In particular, we are able to explain between two to four risk behavior in constrast to the results in Table 4 where we are able to explain between two to six risky behavior.

Taking the share of past lottery expenditure as part of income yields more predictive power because in addition to the WTR (Gen) and the EG (No Loss) the GP task seems to be significant now. Further, taking the log of investment we find improvement in coefficients and its power. Taking the total amount of investment, we find comparable results to Table 4. We find that the HInvQ and the GP task to be significant. This also holds when we take the share of investment made by the household as part of total consumption. The GP task is statistically significant at the 1% level for most of the investment variables while the HInvQ is mostly significant at the 10% level with the exception of the log investment. Hence, variations of the investment variable shows that the significant risk items remain constant.

For the indicator variable of borrowing in general, we found that the WTR (Fin) and the HInvQ was successful in revealing borrowing behavior. This time, however, instead of the WTR (Fin), we find significant relationship with the GP task at the 1% level but also the EG (Loss). In other words, borrowing for business also captures an investment feature which is why both investment questions seem to be significant. Moreover, we find even stronger domain-specific behavior for the number of risk-mitigating activities. The more risk-mitigating activities one decides to implement, the more risk-averse he/she is which are reported by the two investment-related risk items. To conclude with, we can say that despite new variables or an alternative structure thereof, using a single-item risk measure is still unreliable as we succeed in only 17 out of 42 cases.

In the next step we test whether the average of seven items is still significantly superior to the single-item risk measure. In six out of eight cases, our multiple-risk item measure is successful (Table F.1, bottom). Further, for four out of six behavioral items, our multiple-item risk measure is significant at the 1% significance level, indicating its robustness. Regarding factor analysis, again Factor 1 is most relevant, then Factor 2 and finally Factor 3 (Table F.2). The multiple-item risk measure with average of three and two risk items can explain five to six behavioral items (Table F.2, bottom).

Hence, we conclude that our main findings - in particular the robustness of the explanatory power of various multiple-item risk measures over single-item risk measures - also hold for another eight risky behavioral items.

F.2 Role of Cognitive Ability

We examine whether correlation among risk measures is higher for those individuals in our sample with longer education and higher cognitive abilities. Dohmen et al. (2010) measure risk (and time) preferences using a representative sample of 1,000 German adults. They find that people with low cognitive ability are more risk averse. Similar findings are found by Burks et al. (2009) using a sample of subjects in a trucking firm and Benjamin et al. (2013) in a sample of Chilean high school graduates.

These results, however, do not take into account the relationship between risk aversion, cognitive ability, and noise. Andersson et al. (2013) use a representative sample of the Danish population and two standard risk-elicitation tasks - one producing a positive and the other a negative correlation between risk aversion and cognitive ability. They found no significant relation between risk aversion and cognitive ability. Instead, cognitive ability is negatively correlated to the amount of noise. They conclude that errors can cause bias in the estimation of risk aversion from observed choices and that the direction of the bias depends on the specifics of the risk-elicitation task.

Further support that noise is heterogeneous and linked to cognitive ability is provided by Dave et al. (2010). They find that higher math scores are related to less noisy behavior in the multiple-price list task but are unrelated to risk taking. We test the assumption that cognitive ability leads to less noise and therefore improved correlation across elicitation methods. We hypothesize that highly skilled individuals should be making fewer errors and thus produce more consistent results across tasks.

After conducting the risk survey, we asked the respondents six mathematical questions, including addition, multiplication, and percentage calculation. In addition to that, we also tested for word fluency by asking them to verbally list as many animals as they could in 60 seconds (Appendix B.3). The correlation between the two cognitive ability measures is 0.355 (Spearman; $p<0.001$). Thus, the two tests capture a similar underlying trait but also distinct aspects of cognitive ability. We follow the same procedure as Dohmen et al.(2010) and use a single combined measure of cognitive ability using a principal component analysis.

Table F.3 shows that better education and higher cognitive ability improve the correlation coefficients slightly but not dramatically compared to the coefficients for the full sample in Table 4. We find increased correlation between the survey items and between the GP and HInvQ. The largest difference can be seen in the EG experiments, where correlations are overall improved with the other non-incentivized survey items. Essentially, we can infer from the results above that education slightly improves correlation between the experimental measures where probabilities are part of the task. Hence, understanding seems to play some role, yet it is not the decisive factor explaining low correlations. These results partly confirm findings of Andersson et al. (2013) because they also found that noise is task-specific. Therefore, we cannot infer an overall positive direction between cognitive ability and the reduction in noise since we have many different risk elicitation tasks. Thus, we can work with the full sample and do not need to work with a subgroup (i.e. those with higher education or higher average IQ).

F.3 Further Analysis of WTR (Gen)

The responses of our sample population to the WTR (Gen) item indicate an unusually high degree of risk-taking willingness, compared to most other surveys using the same item. Thus we check whether this outcome may distort our findings.

As a first step, we replace the WTR item from the risk survey by the same item from the household panel survey which was conducted only few months earlier (April 2013). It must be noted, however, that we lose observations because we did not always get the same subjects in August. Ultimately, we have information concerning the WTR (Gen) for 512 observations that could be matched. The average response in the household panel survey is 4.51 which compared to the average response in our risk survey of 6.86 comes closer to the findings of Dohmen et al. (2010) and Hardeweg et al. (2013).

Using the WTR item from the spring 2013, we find that it is able to explain households' past expenditure on lottery at the 10% significance level but it is unable to explain further risky behavioral variables (see Table F.4, Panel A); a result also found using the WTR item from August. This result also holds for other analyses, such as using the average of all seven risk items with the WTR item from the household survey; it is still significant in explaining 8 out of 11 items, compared to 9 items before (see Table F.4, Panel A, bottom). We conclude that the high level of risk tolerance in our risk survey does not reduce explanatory power of this item.

In the next step, we compile an average of the WTR item from spring and summer 2013 to see whether results hold in Table 4. Table F.4, Panel B reports results. We find that the average of both WTR items from the household panel survey and risk survey does not change results. The average is able to explain the share of lottery expenditure at the 5% level but no other item which is similar to Table 4.

F.4 Restricting the Sample to the Household Head

Table F.5 replicates the results of Table 4, however, using this time only the household heads as respondents. We find significant changes in results. Household heads seem to be less risk tolerant compared to the spouse who often is a woman. The difference in risk attitude, therefore, goes back to the long-standing debate about gender differences in preferences (Croson and Gneezy, 2009). According to them, one major reason for gender differences in risk-taking is that women differ in their emotional reaction to uncertain situations and this differential emotional reaction results in differences in risk taking. Using household heads reduces the observation to around 400 subjects. Out of these 400 subjects, 303 are male respondents. Moreover, they also are slightly older than the average (59 years old).

Overall, our single-risk items are able to explain 21 out of 77 risky behavioral items which corresponds to the finding in Table 4. We do, however, find significant heterogeneity in terms of the strength of various risk items. While the EG (No Loss) and HInvQ were able to explain five and six behavioral items in Table 4, most of our single-item risk measure in Table F.5 like the WTR (Fin), WTR (Gen), and EG (Loss) are now able to explain up to three items. Hence, it is more equally balanced. Except with the WTR (Gen) in column (7), all significant results have the expected signs. Moreover, it must be noted that the risk-mitigating measure is now explained by four instead of two items. In addition to the HInvQ and the EG (No Loss), we find now that WTR (Fin) and GP are able to explain risk averse behavior in terms of implementing risk-mitigating strategies at the 5% significance level. The same can be observed with health insurance. The EG (Loss) complements the previous two risk items, the WTR (Fin) and HInvQ, in explaining the probability in choosing an additional health insurance at the 1% significance level. For most items, we have an increase in significance. Yet, as a downside, we find that some items are unable to be explained by any single-item risk measure such as the BMI.

Since we hypothesize that the difference in significant results may be due to gender difference, we run the same regression using men only (Results are available upon requests). We find similar results when using male respondents only.

Given the heterogeneity in results, it is important to investigate whether our results still hold using the average of all seven items given the subsample of household heads. We find that our multiple-item risk measure is still superior in explaining behavior (see Table F.5, bottom). It is able to explain six out of eleven items using the household heads only. The average of three risk items is able to predict five behavioral items out of eleven (Results are available upon request). In summary, given the change in subsample, i.e. household heads and reduction in observations, we find that our multiple-item risk measure still has greater external validity and predictive power than using each risk item alone.

Table F.1: Extensions of Risky Behavior

	(1) Lottery (Expend/Income)	(2) Log Investment	(3) Total Investment	(4) Investment (Expend/Cons)	(5) Borrowing (Business)	(6) Number of Risk-Mitigating	(7) Accident Insurance	(8) Over-Weight
WTR (Gen)	0.007**							
	(0.00)							
WTR (Fin)								
HInvQ		0.009***	41.230*	0.016*	0.001***	-0.004*		
		(0.00)	(23.236)	(0.01)	(0.00)	(0.00)		
CEquiv								
EG (Loss)					0.012**		-0.013**	0.026**
					(0.01)		(0.01)	(0.01)
EG (No Loss)	0.016**						-0.016***	
	(0.00)						(0.01)	
GP	0.001*	0.010***	31.929**	0.014***	0.001***	-0.003**		
	(0.00)	(0.00)	(15.68)	(0.01)	(0.00)	(0.00)		
Observations	711	390	715	715	715	715	715	715
Average of 7 Items	0.026***	0.195***		0.244*	0.314***	-0.122***	-0.152*	
	(0.01)	(0.06)		(0.13)	(0.07)	(0.04)	(0.08)	
Observations	709	389		713	713	713	713	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Estimator	OLS	OLS	OLS	OLS	Probit	OLS	Probit	Probit

Notes:
The dependent variables are the behavioral variables from the household survey. Expend/Cons in parenthesis is the total amount of household expenses in the last 12 months as a share of total consumption. Expend/Income in parenthesis is the total amount of household expenses in the last 12 months as a share of total income. See Appendix B for details. Controls include Female, Age, Height, Marital Status, Household Size, Education and Log Per Capita Consumption. We employ the subsample with age of 17-79. Clustered errors on the village level are in parenthesis. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table F.2: Factor Analysis and Multiple-Risk Items (Robustness)

	(1) Lottery (Expend/Income)	(2) Log Investment	(3) Total Investment	(4) Investment (Expend/Cons)	(5) Borrowing (Business)	(6) Number of Risk-Mitigating	(7) Accident Insurance	(8) Over-Weight
Factor 1		0.219***	954.31*	0.378*	0.028***	-0.087*		
		(0.07)	(528.50)	(0.23)	(0.01)	(0.05)		
Factor 2	0.022*				0.022*		-0.034***	0.039*
	(0.01)				(0.01)		(0.01)	(0.02)
Factor 3	0.020*							
	(0.01)							
Observations	709	389	713	713	713	713	713	713
Average of 3 Items	0.029**	0.180***			0.243***	-0.100**	-0.126*	
	(0.01)	(0.07)			(0.07)	(0.04)	(0.07)	
Average of 2 Items		0.191***		0.308*	0.213***	-0.082*	0.223***	0.032*
		(0.07)		(0.19)	(0.08)	(0.04)	(0.08)	(0.02)
Observations	711	390		713	715	715	715	715
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Estimator	OLS	OLS	OLS	OLS	Probit	OLS	Probit	Probit

Notes:
The dependent variables are the behavioral variables from the household survey. Expend/Cons in parenthesis is the total amount of household expenses in the last 12 months as a share of total consumption. Expend/Income in parenthesis is the total amount of household expenses in the last 12 months as a share of total income. Total Investment is the total amount of household investment in the last two years. See Appendix B for details. Controls include Female, Age, Height, Marital Status, Household Size, Education and Log Per Capita Consumption. We employ the subsample with age of 17-79. Clustered errors on the village level are in parenthesis. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table F.3: Spearman's rank correlations for subsample

	WTR (Gen)	WTR (Fin)	HInvQ	CEquiv	EG (Loss)	EG (No Loss)	GP
WTR (Gen)	1.000						
WTR (Fin)	a) 0.381*** b) 0.401***	1.000					
HInvQ	a) 0.128** b) 0.033	a) 0.041 b) 0.092*	1.000				
CEquiv	a) -0.005 b) 0.115	a) 0.070 b) 0.006	a)0.027 b) 0.054	1.000			
EG (Loss)	a) 0.126** b) 0.149***	a) 0.009 b) 0.035	a) 0.031 b) -0.028	a) 0.146** b) 0.161***	1.000		
EG (No Loss)	a) 0.070 b) 0.065	a) 0.015 b) 0.007	a) -0.041 b) -0.048	a) 0.095 b) 0.099*	a) 0.364*** b) 0.405***	1.000	
GP	a) 0.075 b) 0.056	a) 0.045 b) 0.017	a) 0.245*** b) 0.261***	a) 0.014 b) -0.011	a) 0.103* b) 0.125**	a) 0.171*** b) 0.107**	1.000
N: 760							

Notes:
The table reports pairwise Spearman rank correlation coefficients for a subsample with age of 17-79 and
a) Having high education (more than 5 years), (N=288)
b) Having above average cognitive ability, (N=367)
Statistical significance is in parenthesis. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table F.4: WTR Variations and Risky Behavior

	(1) Lottery (Expend/Cons)	(2) Future Lottery Expenditure	(3) Self-employment	(4) Farming	(5) Investment (Expend/Income)	(6) Plan to Invest	(7) Borrowing	(8) Risk Mitigating	(9) Number of Insurance	(10) Health Insurance	(11) BMI
A: WTR Replacement											
WTR (Gen)	0.016*										
	(0.01)										
Average of 7 Items	0.050***	12.893***	0.019*	-0.035*	0.120*	0.046**	0.040**	-0.053**		-0.019**	
	(0.02)	(4.53)	(0.01)	(0.02)	(0.07)	(0.02)	(0.02)	(0.02)		(0.01)	
Observations	477	481	482	485	485	485	485	485	485	485	482
B: WTR Over Time											
WTR (Gen)	0.061**										
	(0.03)										
Observations	477	481	482	485	485	485	485	485	485	485	482
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Estimator	OLS	OLS	Probit	Probit	OLS	Probit	Probit	Probit	OLS	Probit	OLS

Notes:
The dependent variables are the behavioral variables from the household survey. Expend/Cons in parenthesis is the total amount of household expenses in the last 12 months as a share of total consumption. Expend/Income in parenthesis is the total amount of household expenses in the last 12 months as a share of total income. See Appendix B for details. Controls include Female, Age, Height, Marital Status, Household Size, Education and Log Per Capita Consumption. We employ the subsample of 17-79. Clustered errors on the village level are in parenthesis. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table F.5: Only Household Head and Risky Behavior

	(1) Lottery (Expend/Cons)	(2) Future Lottery Expenditure	(3) Self-employment	(4) Farming	(5) Investment (Expend/Income)	(6) Plan to Invest	(7) Borrowing	(8) Risk Mitigating	(9) Number of Insurance	(10) Health Insurance	(11) BMI
WTR (Gen)	0.061**		0.011**				-0.014**				
	(0.01)		(0.01)				(0.01)				
WTR (Fin)							0.019**	-0.017**		-0.006**	
							(0.01)	(0.01)		(0.00)	
HInvQ			0.001**				0.000*	-0.000**		-0.000**	
			(0.00)				(0.00)	(0.00)		(0.00)	
CEquiv				-0.010***		0.009**					
				(0.00)		(0.00)					
EG (Loss)	0.036**								-0.115**	-0.017***	
	(0.01)								(0.05)	(0.01)	
EG (No Loss)		13.944***						-0.039*			
		(5.67)						(0.02)			
GP					0.010***			-0.002**			
					(0.00)			(0.00)			
Observations	392	392	396	394	396	396	396	396	396	396	
Average of 7 Items	0.049*	10.191*	0.031*		0.188**			-0.079**		-0.032**	
	(0.03)	(5.26)	(0.02)		(0.09)			(0.03)		(0.01)	
Observations	391	391	395		395			395		395	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Estimator	OLS	OLS	Probit	Probit	OLS	Probit	Probit	Probit	OLS	Probit	OLS

Notes:
The dependent variables are the behavioral variables from the household survey. Expend/Cons in parenthesis is the total amount of household expenses in the last 12 months as a share of total consumption. Expend/Income in parenthesis is the total amount of household expenses in the last 12 months as a share of total income. See Appendix B for details. Controls include Female, Age, Height, Marital Status, Household Size, Education and Log Per Capita Consumption. We employ the subsample with age of 17-79. Clustered errors on the village level are in parenthesis. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Chapter 3

The Effect of Peer Observation on Consumption Choices: Experimental Evidence

[1]This chapter is based on the article by Grohmann and Sakha (2015). We would like to thank several participants at the IMEBESS and NIBS 2015 Conference for helpful suggestions and especially David Kiss, Stefan Klasen, Sabine Liebenehm, Lukas Menkhoff, Susan Steiner and Juliane Zenker. Financial support by the German Research Foundation (DFG, grant RTG 1723) is gratefully acknowledged.

3.1 Introduction

The feeling of buying something just because someone else has it is a feeling familiar to many. Despite anecdotal evidence that peers exert a very powerful influence over one's consumption behavior, there has been a surprising lack of empirical and experimental research on this topic. In traditional economic theory consumption choices are regarded as a function of budget, price and personal preferences. The effect of those around us is rarely considered. In this study we aim to change this and define peer effects as the simple effect that leads individuals to behave in a similar way to those around them.

The study of peer effects in consumption choices is not just crucial in advancing further understanding of human decision-making, but at a second look, can also have an important effect on policy. Peer effects may influence the success of cash transfer programs, for instance. If peer effects are prevalent they could have an effect on consumption decisions taken with cash grants. At the same time, policy makers who are interested in increasing the uptake of certain good such as health services or innovative technology could use peer effects in order to increase consumption of such goods.

One reason why economists have largely ignored peer effects on consumption choice is that identifying peer effects comes with a number of challenges. Measuring the extent to which peers affect decision making is challenging because social group formation is usually endogenous, complicating causal inference. If belonging to a social group is a matter of deliberate choice, it is difficult to assign causality to the impact of the group itself (Manski, 1993). In addition, since individuals may make simultaneous decisions affecting each other it makes it difficult to determine causal behavior.

Hence, it is extremely difficult to identify peer effects from observational data. We solve this problem by performing an experiment. We use a novel approach and a lab-in-the-field experiment to provide a clear evidence of peer effects on consumption choices in a fully controlled setting where no possible confounding factor can hinder identification. To the best of our knowledge, no experiment of peer effects in consumption decisions has been conducted so far. Our experimental setting allows us to overcome the Manskian problem of contextual effects, correlated effects, and endogenous effects. In this paper, we aim to (i) identify and estimate the existence of peer effects in terms of consumption decisions; (ii) investigate some mechanisms through which peer effects operate; and (iii) evaluate the magnitude of peer influence on certain groups.

We are able to control for personal and local confounding factor because our experimental results can be complemented with a large household survey containing a wide-range of socio-economic information of the respondents and the village in which they live. We perform the experiment in rural Thailand because of the prevalence of close-knit communities. Our respondents live in relatively small villages and have often lived there for many generations. In other words, even though assignment to a group is random; groups are made up of people that actually know each other (Mangyo and Park, 2011).

The design of our experiment is straightforward: we test consumption choices by simply offering respondents the choice between a combination of sweet and salty snacks, i.e. the temptation good[1] (called the tasty treat from here on) and money. The amount of money offered increases in every round by ten Baht whereas the tasty treat stays the same. In the control group, respondents have to make their consumption choices on their own, separated from the rest of the respondents. In the treatment group, each respondent still makes his/her own decision, but all respondents play whilst observing each other. Hence, the only difference in outcome can be attributed to peer observation.

[1]Temptation goods are defined as goods that provide the current self with positive utility, but negative utility to any future self, for instance alcohol, cigarettes or unhealthy foods (Banerjee and Mullainathan, 2010).

Our experimental study has a number of advantages since it tackles problems described by Manski (1993, 2000). One is the problem caused by correlated effects, which states that observed peer effect can be caused by unobserved characteristics that are common to a certain peer group, rather than by the presence of peers themselves. Given our large sample size, we assume that our villages are the same on average and that effects can simply be attributed to peer observation. Randomization of our sample in observing and non-observing groups on the village level combined with detailed information about village and household characteristics also attempts to circumvent the problem of contextual effects. Thus, our research design enables us to directly compare outcomes for those groups that performed the experiment with and without peer observation. In order to identify endogenous peer effects, we use the model of endogenous peer effects with leave-out mean in which individuals average consumption is regressed on the mean of the group average — excluding the individual himself (Angrist, 2014). This econometric strategy allows within-group level variance; circumventing the problem of endogeneity.

We focus in particular on the effect of peer's observation on temptation goods, since consumption choices for temptation goods are particularly susceptible to the influence of peer effects, as has been shown, especially for young people in social psychology (Gunter and Furnham, 1998; Steinberg and Cauffman, 1996). Another reason for choosing temptation goods was that there is no real economic or welfare needs for the temptation goods that are offered. The idea behind this is that playing the game with goods that are necessities may have confounding effects on the demand for the good compared to the money offered.

In order to support our experimental analysis, we develop a theoretical framework. We adjust a standard model of consumer choice with a cost imposed on the decision maker when deviating from the group choice. We argue that this cost represents a social cost from not conforming to the group. We can show from this model that under the peer treatment extreme choices are more costly and therefore, the demand curve for the tasty treat is flatter under peer treatment.

Our experimental data confirms the prediction of the model. Specifically, we find that observing groups - those that sit in close proximity with each other - have a higher group minimum and a lower group maximum. Consequentially, the standard deviation for those observing groups is lower than for those groups that simply played at the same time, but without peer observation.

In further analysis, we confirm this finding by showing that the group average, excluding the individual him/herself, has a positive and significant influence on the decisions made by the individual respondent; however only when the experiment is performed with peer observation. Most importantly, the effect is not significant when the experiment is performed in non-observing groups. The effect being only present in the observing groups shows us that it can be attributed to the presence of peers and not to other possible effects.

Next, we aim to explore the mechanisms behind the peer effect. There are two possible reasons for this; either the respondents feel that the others in a group have better information or they are gaining some kind of benefit from conforming to others. We find evidence that unfamiliarity with a product is counteracted by peer observations, indicating some evidence for the first mechanism.

Subsequently, we look at treatment heterogeneities to analyze whether there is a different magnitude of peer effects for individuals with different background characteristics. We show that those with the highest cognitive ability are less susceptible to peer effects. Using the same technique, we do not find any effect for overconfident, underconfident or higher income respondents.

To sum up, we are able to show using a lab-in-the-field experiment that the observation of peers has a significant impact on consumption choices. We find evidence of convergence in consumption choices when observing one's peers. Our results contribute to the literature on conformity and herding behavior where conformity is defined as an intrinsic taste to follow others (Goeree and Yariv, 2010), driven by factors such as popularity, observational learning, esteem and respect (Bernheim, 1994).

A number of experimental studies use a similar experimental design; Falk and Ichio (2006), randomly assign participants either to a group or work alone in order to study the effect of peers on productivity. In another study by Baecker and Mechtel (2014) use a similar design in order to study the effect of peers on cheating behaviour. These studies have the advantage that they provide a clear counterfactual and control for local factors, thus providing the cleanest evidence on peer effects.

Further empirical evidence on peer-group phenomena can be found in the context of other economic behaviors [2]. These results, however, rely mostly on observational data which makes it difficult to separate peer effects from the effect of confounding factors.

Recent papers try to measure peer effects in terms of education using, for instance, natural experiments to overcome the reflection problem such as the random assignment of college students to their respective dorms (Sacerdote, 2001) or the exogenous influx of students in neighboring schools after the hurricane Katrina (Imberman et al., 2012). However, even with 'natural' experiments when the setting offers an enhanced strategy to identify peer effects, the impossibility of controlling for all local or personal confounding factors does not provide a clean identification strategy.

Another recent strand of literature uses the existence of partially overlapping groups of peers to solve issues related to both reflection and correlated effects. The intuition is that partially overlapping groups generate peers of peers (or excluded peers) who act as exclusion restrictions in the simultaneous equation model of so-

[2]Peer effects have been studied in the context of other economic behaviour. Peer effects seem to have a positive impact, for instance, in terms of workers' productivity (Guryan et al., 2009; Mas and Moretti, 2009; Bandiera et al., 2010), education (Sacerdote, 2001; Imberman et al., 2012), technology adaptation (Munshi , 2004; Oster and Thornton , 2012), and saving and investment decisions (Duflo and Saez, 2003; Viscusi et al., 2011).

cial interactions and, thus, solve the reflection problem.

Something that has been rarely attempted so far when looking at peer effects is to distinguish between the different reasons that cause individuals to behave in a similar way to their peers. To our knowledge, this distinction has so far only been attempted by using carefully designed field experiments. Cai et al. (2009) look at an experiment with two treatments in a restaurant setting in order to distinguish the effect of social learning from the effect of salience. Burszytyn et al. (2012) study the demand for a complex financial fund, using a brokerage firm in Brazil. The author's aim to distinguish between wanting what others have and the information effect of knowing what the other person thinks.

The remainder of this article is organized as follows. In Section 3.2, we present our data and experimental design. Section 3.3 presents the conceptual framework and the identification strategy. We discuss descriptive statistics, inferences and results in Section 3.4. Section 3.5 provides further robustness tests while Section 3.6 concludes.

3.2 Data

3.2.1 Household Survey Data and Sampling

Our peer experiment was conducted as part of a larger household survey of the research project "Impact of shocks on the vulnerability to poverty: Consequences for development of emerging Southeast Asian economies" funded by the German Research Foundation which was conducted in three Northeastern provinces of Thailand since 2007. The household survey contains detailed information on many aspects of households' living standards including: household demographics, recurrent and durable expenditures, credit and savings, landholdings, agriculture, employment, health, and education. It also includes materials concerning village characteristics such as the number of village institutions or infrastructure (i.e. irrigation system, access to electricity, nursery etc.), in – and outward village migration, inhabitants, but also the number of shocks occurring in a village. This data provides a representative sample of rural households in the Northeastern part of Thailand.

Our peer experiment was conducted in the largest of the three provinces, Ubon Ratchathani where the main source of livelihood is subsistence agriculture, and seasonal labor work. In addition to standard socio-demographic variables from the household survey, we also collected a number of variables that are designed to measure cognitive abilities after the peer experiment. We collected two types of questions (Details are reported in Appendix B. Firstly, we collected a number of math based questions. In total there were six questions, the first four are based on the hardest four out of eight math questions in Cole et al. (2011), the last two questions are based on question used in the Survey of Health, Ageing and Retirement in Europe (SHARE). In addition, we included a question that asks respondents to name as many animals as they can in 60 seconds. This is a measure of word fluency and has the advantage that it is related to more innate forms of intelligence and especially measures processing speed. This test for word fluency has also been used in a number of other studies as part of cognitive ability measures such as Falk et al. (2010).

Finally, we ask respondents to judge how many of these questions they answered correctly to measure overconfidence. Overconfidence results in unrealistically positive self-evaluations. In other words, people are unrealistically optimistic and overestimate personal success probabilities but may also have a status and signaling effect. Our primary measure of confidence is the difference between the predicted math score and the achieved score. Thus, a subject whose prediction is higher than her actual score is called overconfident, and a subject whose prediction is below her actual score is called underconfident.

The sampling procedure of rural households for the peer experiment conducted in Ubon Ratchathani follows a three-stage stratified sampling procedure. It is important to know that we exclude the urban area around the provincial capital city and confine the sample to the remaining rural areas. In the first stage subdistricts within the province were chosen with probability proportional to size and implicit stratification by population density. In the second stage, from each sampled sub-district, two villages were sampled randomly with probability of selection proportional to size. In the last step, in each of those villages a systematic random sample of ten households was drawn to be interviewed from the household lists of the rural census ordered by household size. To conclude with, villages as well as respondents were randomly sampled for our peer experiment.

3.2.2 Experimental Design

The peer experiment was conducted in August 2013 with a total of 521 respondents from 66 villages in Ubon Ratchathani. We link experimental results with the large household survey that provides us with individual-level demographic information.

The experiment was carried out by local enumerators with one of the co-authors being present at all times. Instructions were translated from English into Thai and back, and were cross-checked by a Thai economics professor to avoid semantic difficulties. Instructions were kept as simple as possible Appendix A. The interviewers were trained in sessions that lasted a total of five days. During these five days, a pilot study was conducted in three villages.

We randomly assigned the villages to their respective treatments. The experiment was conducted by visiting two villages per day; one in the morning and one in the afternoon. For neighboring villages experiments were usually carried out simultaneously. The distance between villages was on average 18 km and respondents had to stay at the experimental site until the completion of the survey. There were two experimental sessions conducted in each village, with up to five respondents in one session at the same time. All experimental sessions took place in the village hall.

The experiment consists of a very simple choice task that required no previous knowledge, was easy to implement and to measure in the field with the rural sample (Appendix A). The respondent has to choose between the tasty treat and a certain amount of money. Before the experiment, however, respondents were asked to estimate the price of the tasty treat. After their prediction, all respondents were told that the tasty treat costs 40 Baht in order to avoid information asymmetries concerning the value of the product. Another important component of our pre-experiment data collection was that they could receive the tasty treat right after the experiment while money at the end of the survey, thus enhancing temptation. Time-discounting factors can largely be ignored since the experiment, including post-experiment questions, only took between one to two hours to complete. Respondents were also reminded that they had to stay and answer further questions (risk attitude, financial literacy, overconfidence and cognitive ability) after the experiment.

The tasty treat consists of very popular items that are widely known across the country - a can of coke, a piece of custard cake, a small package of lays classic crisps, a bar of chocolate, and a small pack of candies. It had a value of 40 Thai Baht (THB) (approximately 1 Euro). We made sure that it not only included sweet but also savory items so that it appeals to a wider range of tastes. During the experiment, we made sure that the respondents did not get any food or sweet beverages to drink. Furthermore, we made the tasty treat visible for the respondents whilst playing the choice task in order to increase temptation.

In the choice task, the respondent has to choose between the tasty treat and a certain amount of money. In total, every respondent has to make this decision seven times. In the first round, the respondent has to choose between the tasty treat and 10 THB. Once the respondent makes the decision, the amount of money is increased by 10 THB. Since we have seven rounds, in the last round the respondents have to choose between the tasty treat and 70 THB. In round four there is no price difference between the two choices. After round four, it becomes increasingly unreasonable to choose the tasty treat because of the significant price difference. The enumerator marks the decision in each round. We did allow switching back and forth. There were 24 respondents who switched twice and were dropped later in our sensitivity analysis.

Once all seven choices have been made, one decision was randomly played out by picking a number from a non-transparent bag that goes from 1 to 7. In case the respondent picked number 3 and choose the tasty treat in row 3, she received the tasty treat immediately. In case, the respondent picked money in that row, the respondent would receive 30 Baht at the end of the survey with an additional 50 Baht for participating in the survey. After the experiment, respondents were asked how much he/she would be willing to pay at most to receive the tasty treat.

In the control group, the tasty treat game was played individually and was conducted with 261 individuals in 60 groups. To avoid peer observation, we made sure that respondents were separated across the town hall so that they could neither hear nor see the choices of the other respondents. Furthermore, it is unlikely that the decision of one respondent affects other respondent in the control group because individuals respond at different speed levels.

The peer treatment was conducted with 260 individuals in 66 groups. The size of the group ranges from three to five people. The procedure of the treatment is the same as the individual treatment with the sole exception that decisions were conducted with peer observation. Each respondent is still responsible for their own decision, but they have to sit next to each other and perform the experiment. As in the control group, all the instructions were read out loud and show cards were used to demonstrate the choices between tasty treat and money in each round. In the first round, for instance, the principal enumerator asks all respondents whether they would like to choose the tasty treat or 10 THB. Respondents have to express their choices to their assigned enumerator out loud so that other participants could hear and see their choices. Once everyone has decided in each round, the principal enumerator moves to the second round and the same procedure follows.

Given our experimental design, we cannot observe an order in which participants answer. What we can observe, however, is whether the spatial and social proximity of the peer around the table in conjunction with their announcement of the decision into the group affects consumption choices of individual. The difference between the treatment and the control group is simply that choices are observable to peers.

3.3 Conceptual Framework and Identification

3.3.1 Conceptual Framework

In this section we present our conceptual framework that explores the relationship between the choice of money m, the choice of a tasty treat tt and the group's choice of $\bar{tt}$. In this section, we ignore the effect of individual preferences as denoted by x and $\bar{x}$ in the next section. We can justify this with our experiment and due to personal preferences being the same across treatments. Hence each participant's utility function is defined as:

$$U(tt, m; D, \bar{tt}) = u(tt, m) - D \cdot c(tt - \bar{tt})$$

The first component $u(tt, m)$ is both increasing and concave in both tt and m. It represents the utility that an individual receives from choosing the tt or m, whereas the choice in $tt \in \{0, 1\}$ and $m \in \{10, ..., 70\}$. Because individuals have to decide between tt and m, $tt = 1$ implies $m = 0$ and $m > 0$ implies $tt = 0$. Also note that the difference $u(0, m) - u(1, 0)$ is increasing in m: the higher m, the smaller the share of individuals that will prefer tasty treatment to money, i.e.

$$\frac{\partial \Pr(tt \succ m|D)}{\partial m} < 0.$$

The utility function above includes a conformity cost function $c(tt - \bar{tt}) \geq 0$. This cost function is increasing, the larger the difference between own choice of the respondent and average consumption of the peers.

$$c(tt - \bar{tt}) \begin{cases} > 0 & \text{if } tt \neq \bar{tt} \\ = 0 & \text{if } tt = \bar{tt} \end{cases}$$

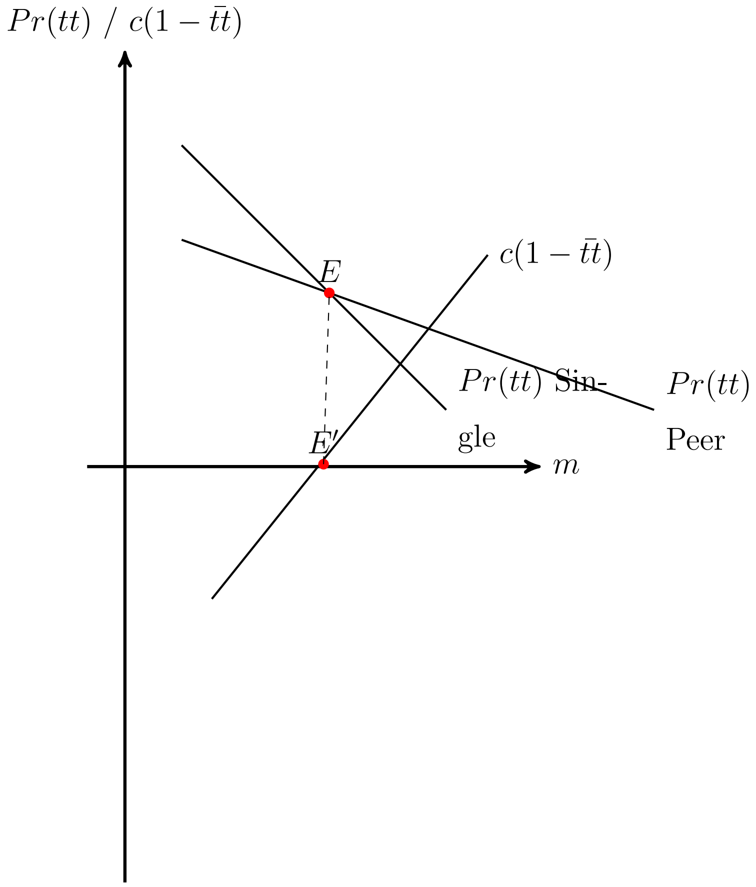
In this model we do not go into the source of this cost. In our view there could be a number of reasons behind this, which we discuss later on. More importantly note that this conformity cost only applies to those individuals that play in a group. In this case $D = 1$ and for individuals in the control group $D = 0$ hence the conformity cost function does not play a role in their decision making. In single treatment, the tt is preferred if

$$u(1,0) > u(0,m).$$

In the group treatment, tt is chosen if

$$u(1,0) - c(1 - \bar{tt}) > u(0,m) - c(0 - \bar{tt}).$$

As peers possess the same utility function $U()$, average peer tasty treat consumption $\bar{tt}$ must also be decreasing in m. Therefore, $\frac{c(1-\bar{tt})}{\partial m} > 0$ and $\frac{c(0-\bar{tt})}{\partial m} < 0$. It should be noted that $\bar{tt}$ also depends on tt and is therefore endogenous. Since choosing the tt is synonymous with not choosing m, it is easier to think of one cost function that looks at the cost of choosing tt at different levels of m. In this case the cost of choosing tt would be positive for high values of m, but negative for low m. **Figure 1** shows the relationship between m, $Pr(tt)$ and $c(1 - \bar{tt})$.



Thus, it becomes clear that the respondents under peer treatment react more strongly to a change in m than respondents under single treatment

$$\frac{\partial \Pr(tt \succ m|D=0)}{\partial m} < \frac{\partial \Pr(tt \succ m|D=1)}{\partial m}.$$

Intuitively, this seem logical as there in an extra benefit from choosing the tt when m is small and an extra cost in choosing tt when m is large. This means that hence in the peer treatment, we expect that fewer people switch from m to tt at an early or late stage. In turn we expect this to lower standard deviation within a group. So far we have shown the different reactions of tt to a change in m, between the peer and the single treatment. We are now left to show that the aforementioned conformity cost leads to a positive relationship between tt and $\bar{t}t$ which can be defined as peer effects.From the original utility function we can see that

$$\frac{\partial \Pr(tt \succ m | D = 0)}{\partial \bar{tt}} = 0$$

Hence there is no change in tt as $\bar{tt}$ change in the single treatment. Whereas under peer treatment

$$\frac{\partial \Pr(tt \succ m | D = 1)}{\partial \bar{tt}} > 0$$

There is a positive relationship between the number of people that choose tt and the average peer decision $\bar{tt}$.

3.3.2 Identification Strategy

We are interested in identifying causal peer effects and understanding whether and how much consumption is affected by the observation of peers. The identification of peer effects, however, suffers from a number of econometric issues (Manski 1993, Moffit 2001) which can be summarized into three categories: (a) contextual effects, (b) endogenous effects, and (c) correlated effects.

Contextual effects in consumption choices may emerge if socially-related individuals under study share preferences and characteristics that make them more likely to select in a peer and these characteristics are important determinants of the dependent variable. Correlated effects may emerge if individuals share common environments and unobserved shocks (i.e. rainfall in the village) that make their consumption move simultaneously independently of any genuine peer effects. Finally, endogenous effects represent the phenomenon where the group affects individual behavior through social interaction (i.e. is the individual's consumption choice positively or negatively affected by the group consumption choice?). It is the third effect what we are trying to separate in this study.

Our experimental design (discussed in detail in Section 3.2) represents an attempt to surmount the challenge of identifying a causal peer effect. Much of the literature following Manski has focused on the econometric issue of separating the causal peer effect from that of correlated unobservables (Conley and Udry, 2010; Miguel and Kremer, 2004; Bandiera and Rasul, 2006). Two ways of disentangling these effects are to (1) randomize the peers (Sacerdote, 2001; Duflo and Saez, 2003) or (2) randomize an intervention or new technology (Oster and Thornton, 2012; Godlonton and Thornton, 2012; Kremer and Miguel, 2007). We follow the first approach.

The double randomization in our experimental design, that is, first randomly selecting households to perform the experiment given the sampling procedure and second randomizing peer and control treatments according to villages, circumvents the problem of correlated and contextual effects. Given our random assignment of individuals to play the game alone or in a group, we are able to create counterfactual groups out of those individuals that played the game at the same time as their peers – without directly observing their peers. We have two types of groups, those that performed the experiment directly observing each other and those that played the game at the same time in the same room, but not directly observing each other. Hence the only difference between our treatment and control group is that the treatment group observed their peers and the control group did not.

As a check of the randomization, in Table 2 and Table 3 we present individual's characteristics for the observing and non-observing groups, as well as tests of equality of characteristics across groups. As expected from the random assignment into each group, the sample is well balanced across the baseline variables. We try to overcome the reflection problem by the identification of endogenous peer effects with the so-called leave-out mean which we use as the regressor in our main analysis to analyze the effect of the group average consumption on the individual consumption choice.

To identify the effect of peer observation, we will estimate the main regression model in the following form using least squares estimation:

$$Y_{i,j} = \alpha T_j + \beta \bar{y}_{-i,j} + \gamma \bar{x}_{-i,j} + \delta x_{i,j} + u_{i,j}$$

In our framework, $Y_{i,j}$ is the consumption choice of tasty treat for individual i who has group affiliation j (observing or non-observing group). In our main analysis $Y_{i,j}$ will be the last row in which they choose the tasty treat before switching to money. However, we also run similar regressions using an indicator variable if they always chose tasty treat over money or if they decided not to choose tasty treat at all. The T_j is a dummy equal to 1 if the group j was assigned to the peer treatment. The coefficient of interest β is the sample mean of the group outcomes, net of individual i's outcome, a quantity commonly referred to as the leave-out mean.

In many peer studies, often researchers would often use the group mean inclusive of the individual, $\bar{y}_{ij}$. However, outcome-on-outcome peer effects are vacuous, because regressing $\bar{y}_{ij}$ on y_{ij} results in a coefficient of 1, entering unity. Therefore, any peer group measure must vary within groups in order to satisfy the rank condition. This would rule out taking the average outcome of the group as the regressor. Instead taking the leave-out mean allows inter-group correlation coefficients since there is a different group average for each respondent, calculated from the decision of the other group members. This approach has previously been used by Townsend (1994), Guryan et al. (2008), Duflo et al. (2011), Carrell et al. (2012) and advocated by Angrist (2014).

Following this, we include the variable $\bar{x}_{-i,j}$ which is the vector of average individual's socio-economic characteristics in group j, excluding the individual i. A set of individual characteristics such as female, age, schooling, log consumption, household size, dependency ratio, algebra knowledge and BMI that affect consumption decisions compose $x_{i,j}$. In the robustness section, we also include specific village characteristics such as the travel distance to the district capital, the provincial capital, the average number of shocks a village experienced in the last two years and the number of households living in the village.

The error term u_{ij} is clustered on the village level. Following the literature, $\bar{y}_{-i,j}$ measures the endogenous effect, If $\beta > 0$, positive peer effects persist in a group, $\beta = 0$ implies absence of peer effects, while these effects are negative if $\beta < 0$. The parameter $\delta x_{i,j}$ is the exogenous effect, $u_{i,j}$ is the correlated effect.

As we assign respondents randomly into peers groups, we assume $E(u_{ij}|x_{ij}) = 0$, i.e., no correlated effects or self-selection into groups. In our particular case, the randomization of individuals into observing and non-observing groups rules out correlation between the individual effect and any endogenous or exogenous effect, thus satisfying the condition, $Cov(E(\bar{y}_{-ij}|u_{ij}) \neq 0$. In other words, since u_{ij} is not mechanically correlated with $\bar{Y}_{-ij}$, we can avoid the classical simultaneity problem and infer a causal relationship. Thus, if we observe a difference in outcomes between observing and non-observing groups we can attribute this directly to the (on average) only difference between these groups, namely peer observation.

3.4 Results

3.4.1 Descriptive Statistics

Table 1 shows individual characteristics of our sample. First, we have significantly more women in our sample (60%). As we are deliberately sampling the household head, average age is relatively high at 54 years and 82% of respondents are married. Socio-demographic characteristics of our sample are typical for rural northern Thailand; education levels are still relatively low with less than six years on average. The average household has more than four members with a dependency ratio of 1.45 dependents for every working member. The vast majority of respondents name farming as their main occupation, with the rest being made up of government officials, business owners, students and housewives. As this study uses eatable goods to examine the consumption of temptation goods, it is interesting to look at BMI, a standardized measure of weight to height ratio. The average in our sample is 23 which is the normal BMI range according to the WHO. In terms of village characteristics, the average distance to the enxt district capital and to the provincial capital, Ubon, is 16 km and 60 km respectively. This is important to know and to control for because the demand for the temptation good may be larger the higher distance of the village to the nearest town. The average number of shocks in our 66 villages was 1.45 ranging from 1 to 3 shocks in total. The number of households in a village varies significantly from 813 households close to the provincial capital to 55 households only which has also the highest distance to the provincial capital Ubon. Peer effects may be larger the smaller the village is because people may know each other better. Despite considerable growth in rural Thailand over the last decades, the north east is still relatively poor which is reflected in the average rate of consumption and average household wealth.

In addition to standard socio-demographic variables, we also collected a number of variables that are designed to measure cognitive abilities. This allows us to study the peer effect on a sample with different levels of cognitive ability. Firstly, we collected a number of math based questions. In a first step, we awarded one point for each question answered correctly. The average score achieved is 3.6 out of six. Numeracy shows a near normal distribution with 1.99% scoring no point and only 4.81% scoring full six points. Second, we asked respondents to name as many animals as they can in 60 seconds. The average number of animals named is 17.29; however the standard deviation for this measure is rather large at 5.86. The correlation between the two cognitive ability measures, numeracy test and word fluency is 0.355 (Spearman; p-value<0.001). Thus, the two tests capture a similar underlying trait but also distinct aspects of cognitive ability. Third, we follow the same procedure as Dohmen et al. (2010) and use a single combined measure of cognitive ability.

Finally, we also measure overconfidence of our respondents to see whether over/underconfident respondents are more resisting or susceptible to peer effects. We define a subject whose math prediction is higher than his/her actual score as overconfident, and a subject whose prediction is below her actual score is called underconfident. Using this measure, 40% of our sample are overconfident while 33% are underconfident. We find a positive correlation between cognitive ability and over/underconfidence. Overconfidence is positively correlated with lowest 10% of cognitive tests (0.26, p-value<0.001). In contrast, the correlation between high cognitive skills (highest 10%) and underconfidence are 0.09 with a p-value of 0.001.

Table 2 shows results of our paired T-test and Wilcoxon rank-sum test to check for differences between treatment and control groups. This shows that randomization was mostly successful and that there are no significant differences in observables between those that played the tasty treat game alone and those that played the game with peer observation. The only difference that can be seen is that those played in a group have on average more children which is statistically significant in the T-Test and in the Wilcoxon rank-sum test. The distance to Ubon, the provincial capital, is also larger in the control group, according to both tests. We will hence control for this difference in further analysis.

As this study not only compares the behavior of individuals but also looks at the behavior of groups, it is important to check that group composition is the same between those that played in observing and non-observing groups. There are 126 groups in total. 60 groups are observing group, while the rest played are non-observing group. Table 3 shows that group composition stays mostly the same on average when looking at measured observables. In line with Table 2, Table 3 shows that on average respondents assigned to the treatment group have a higher number of children, which is significantly different for both the t-test and the Wilcoxon –rank-sum test. We control for possible confounding effects in later regressions.

3.4.2 Comparing Groups

We begin our analysis of the effects of peer observation by studying the difference between those groups that played the game observing each other and those that played the game at the same time and under the same conditions but not observing each other. T-tests and Wilcoxon-rank test compare decisions between the two types of groups (see Table 4). At first, it seems that there is no difference in the mean of the last row that was chosen in each playing situation. Hence, the average last row chosen in observing as well as non-observing groups is the same. However, we can see a difference in the standard deviation between those groups that played together and those that did not. The standard deviation within a group for those groups that observed each other is significantly lower than for those groups that did not observe each other. Those that play in a group are less likely to switch either very early or very late. This can also be seen when looking at the group minimum and the group maximum. The group minimum is the lowest switching point within the group, whereas the group maximum is the highest switching point within a group. We can see that the group minimum is significantly higher and the group maximum is significantly lower when the game is played with peers observing each other. In other words, we find evidence for converging consumption choices when respondents observe each other. This finding is in line with Falk and Ichino (2006) who find that the standard deviation of output from subjects who have been allocated in pairs is statistically significant different from those in the single treatment.

We further test the finding described above, using regression analysis with results shown in Table 5. Outcome variables stay the same as above, namely group mean, group minimum, group maximum, and group standard deviations. The peer dummy is unity if the group played with peer observation. In these regressions we control for group level characteristics. We confirm our finding from above. When the experiment is played with peer observation, standard deviation of choices within the group is lower. The same can be seen when looking at the group minimum and maximum. The coefficient on the peer dummy is positive in the regression estimating the group minimum and negative and significant in the regression estimating the group maximum. Interestingly, group composition seems to have only a limited influence on the tasty treat choice. Groups with more women switch from tasty treat to money earlier. Similarly, there seems to be an effect of groups that are richer, i.e. that have higher average consumption. Both Table 4 and Table 5 show that there is a significant difference in consumption choices between observing and non-observing groups. We find that consumption choices converge in the observing groups.

This result is in line with theoretical models of herding (Banerjee, 1992; Bikhchandani et al., 1992). Experimental evidence for conformity is found for example in Cai et al. (2009) where customers of a Chinese restaurant learn from the information contained in the choices of others and behave in the same way. It also is in line with the experimental finding of Bolton and Ockenfels (2000) and Fehr and

Schmidt (1999) who find possible explanations for herding behavior such as the reduction in expected inequality or inequality aversion among subjects.

3.4.3 Peer Effects

As a next step we look at peer effects in detail. We are approaching our main result, namely that we find an effect of the group average on the individual decision-making concerning consumption choice. We investigate an endogenous peer effect calculating the group average by excluding the individual him/herself and employing the empirical strategy as discussed in detail in section 3.3.2. This means to calculate a different group average for each respondent. This approach has previously been used by Townsend (1994). Results are presented in Table 6, Panel A. In the first two columns, we take into account inconsistent individuals who switch back and forth in the choice task. We find that there is a significant and positive relationship between the average switching point in the group and individual's switching point. The first two rows cover the entire sample. When looking at these results only, it is conceivable that this relationship may be caused by unobserved variables as described by Manski (1993).

In the next four columns, however, we split out sample into those that played the game in observing groups and those that played in non-observing groups (denoted as Peer and Single). Here, we can clearly see that the effect observed above is caused by peers observing each other directly and not caused by unobserved correlated variables. In columns 3 and 4 we show results for those individuals that played in observing groups with and without control variables. In other words, a one-unit increase in the average switching row for the group from the tasty treat to money leads to an increase in the individual's switching point by 0.63 controlling for group average and individual characteristics. We can see that the effect here is significant and stronger than for the full subject pool. The only significant relationship we find is the difference between males and females. Females seem to choose less the tasty treat compared to money. Columns 5 and 6 show the same regression but for respondents that play the game without peer observation. Here the effect of the average peer choice has no effect on the individual's switching row. Similarly, in column 7 we introduce an interaction term between the group average and a dummy that takes the value one if the game was played in an observing group. The interaction term is positive and significant and so we can conclude that the relationship between the group average and the point of switching is not the same between observing and unobserving groups.

These results in Table 6 Panel A, described above indicate to us that the peer effects that we observe in columns 3 and 4 above, is not caused by unobserved correlated variables but rather by peers observing each other and making the same observation at the same time.

In Table 6, Panel B and Table 6, Panel C we perform the same exercise, but with different dependent variables to see whether individuals consumption decision is still affected by the group. In Panel B we use a dependent term that is a dummy taking the value of 1 if the respondent chose the tasty treat in every round. In Panel C we also use an indicator variable that is 1 if the respondent never chose the tasty treat, hence prefering the money from row 1 until row 7. Both tables exhibit the same pattern as the previous table. The group's average consumption choice does influence the individual's choice in both tables. In Panel B, for instance, it can be said that if the group consumption average increases in the peer treatment, the more likely is the respondent to choose the tasty treat in every round by 0.47 percentage points. It is highly significant at the 1% significance level controlling for observable factors. Noticeable is also the relative high R^2. The baseline covariates plus the group average excluding the individual seem to explain a significant share of the variation in the dependent variable, that is choosing always the tasty treat. Conversely in Table C, we find that if the group average is higher, it is less likely the respondent never chooses the tasty treat. Most importantly, this relationship only holds if the decisions are made under peer observation and does not hold if the game is played at the same time but without observation. Hence, we find that observability of the behavior of peers matters to determine conforming behavior.

These results corroborate the social norm mechanism that individuals feel bad about deviating from the average consumption of their peers. More observable behaviors are largely subject to peer effects than non - observable behaviors. The direction of our effects is in line with those of Bandiera et al. (2010), Bursztyn et al. (2014) and de Giorgi (2010) for positive and significant peer effects on individual behavior.

3.4.4 Mechanisms

So far, we have shown that the standard deviation of choices in groups is smaller, as the maximum switching row in groups is lower and the minimum switching point is higher if the experiment is conducted with peer observation. At the same time, we were able to show that individuals are clearly influenced by their groups, as group averages have an influence on the individual decision.

In this section, we will now attempt to look into the mechanism that operates these observed effects further. In the literature, a number of reasons behind peer effects are discussed (Cai et al., 2009, Bikhchandani et al., 1998). We here attempt to look at two factors. Firstly, peers effects have been argued to be caused because respondents believe that others have better information. Secondly, individuals could simply follow their peers because they are gaining some kind of network externality from doing the same as others in their group. Due to the set up of our experiment, we are unable to provide definite answers. None the less, these results provide some interesting insights into the mechanisms that are behind out observed peer effects.

We will here discuss the effect of information since peer effects have been extensively examined in the context of informational spillovers (Kohler et al., 2001;

Conley and Udry, 2010; Miguel and Kremer, 2004; Foster and Rosenzweig, 1995; Oster and Thornton, 2012). As described above, we asked respondent to estimate how much the tasty treat costs to buy in a shop. We use this response as a proxy for how familiar the respondents are with the product. We create a dummy that is unity if the respondent wrongly estimates the price. We introduce this dummy, together with an interaction term between the dummy and the leave-out-mean into the regression as described above. Results are shown in Table 7. Interestingly, unfamiliarity with the tasty treat makes the respondent less likely to choose it, but only in the single treatment. Not knowing the price of product has no effect on the choice likelihood to choose the tasty treat in the peer treatment.

These results indicate to us that peer observations counteract the effect of a lack of information on a product. Gaining information from peers, therefore, seems to play a role in peer effects. At the same time, we find evidence of people following each other. However, we cannot draw definite conclusions about the mechanism behind peer effects. Network externalities could be at play here in addition to information effects.

3.4.5 Treatment Heterogeneity

In this section, we test whether certain personality types are more likely to succumb to peer effects. It is conceivable that high (low) cognitive ability individuals within their group are able to resist (succumb) to peer effects. To the best of our knowledge, this is so far the first study looking into within-peer dynamics. Yet, there is a growing literature linking cognitive ability and financial literacy to improved financial behaviors and outcomes (see for instance Agrawal and Mazumder, 2013; Bertrand and Morse, 2011). We hypothesize that high cognitive ability individuals should be less subject to peer pressure while the opposite should be true for low skilled respondents.

As discussed above we included a number of question designed to study cognitive ability. We created a dummy for those that score the highest and lowest cognitive ability score compared to their peers within the group. As before standard errors are adjusted for clustering at the village level. Table 8 shows results. We find that the high cognitive ability individuals are less likely to choose the tasty treat at higher rows. The result is statistically significant (p-value<10%) and robust even after we control for socioeconomic characteristics. We, however, do not find statistically significant results for low skilled people.

In the next part, we would like to investigate whether overconfidence and intelligence drives economic decision in our peer experiment. We hypothesize that those who are overconfident may drive the group's decision towards taking more tasty treat.

Similar to the procedure with cognitive ability, we created a dummy for those in the group who are overconfident. While we do get the expected sign, that is to say, that overconfident people take more tasty treat, the results are statistically insignificant. Hence, we do not find that overconfidence matters in terms of peer decisions (Details upon request).

The same analysis as in Table 8 has been performed for higher and lower consumption individuals within the groups. We generate a dummy for the highest and lowest consumption of individuals compared to their peers in the group and do not find any effect (Details upon request). We also use the interaction between cognitive ability and income but results remain unchanged. The only influence we find is that those that performed better in the math test are more able to resist the peer effects in choosing the temptation good.

3.5 Robustness and Discussion

Strictly speaking it is conceivable that the peer effects that we observe in section 3.4.2 are not caused by peer effects since our randomization took place on the village level rather than the individual level, however this seems very unlikely. However, for this to happen the randomization would had to work in a way that those that played in observing groups are more alike than those that played in non observing groups. Since the number of groups is fairly large and we are also able to control for a large number of observable factors, we believe that such a concern can be neglected. However, to further exclude doubt, we test whether standard deviations of observed variables are the same between observing and non-observing groups. Results are shown in Table 9. From the T-test and Wilcoxon-rank test, we can see that standard deviations are the same for observing and non-observing groups. We therefore reject the idea that our results are caused by observing groups being more similar to non-observing groups.

Furthermore, we check whether in addition to the distance of the villages to Ubon, the provincial capital city, or the nearest district capital has an impact on the demand for temptation goods since it is assumed that villages that are close to urban areas could get the tasty treat more easily. This could determine the impact of peer effect. We do not find that the distance to the provincial capital or the district capital has any impact on the peer effects and results found in Table 6 stay the same.

Furthermore, we check if there is an effect of higher food consumption on the likelihood of choosing the tasty treat. We find no effect of food consumption. Next, we also check if the main results hold when we change the way the dependent variable is coded. In order to do this, we create two dummies. The first takes the value of one if the respondent switched before the money amount increased to 40 Baht, the second takes the value of one if the respondent switches after the money amount is increased to 40 Baht. We run all the regressions again and find that the results do not change. The group average still has a significant effect on these outcome variables.

In the next step, we investigate whether morning or afternoon sessions would have a confounding effect on the demand of consumption good (see Table 10). We create an indicator variable for the morning and afternoon session and interaction terms thereof with the group average excluding the individual and include this in our regression analysis described in section 3.3.2. We find that there is no difference between results played in the morning or afternoon sessions. In the first column – including a simple dummy in the peer treatment – we find that group choice with regard to the tasty treat still influences the individual at the highest significance level. In the single treatment, we still find no statistically significant results. This result is confirmed in the last row where we include the interaction between the peer treatment and the morning sessions. We find that there is a statistically significant difference between the average consumption choices between both treatments. Hence, main results of Table 6, Panel A, remain unchanged.

In the next step, it is interesting to see whether the morning sessions matter for the demand of the tasty treat. In each village, we played two sessions. We focus particularly on the sessions with the peer treatment since it seems that the individual is affected by the group choice. We find that the first and the second session does not make a difference in results (Table 11). Taking the entire or merely the observing groups, we find that regardless whether one group played before the other, peer observation seems to have an impact on the consumption choice of the individual.

Further, it would be interesting to investigate whether peer effects are stronger for some subsamples. We first look whether villages with fewer households have a stronger magnitude of peer observation. We, therefore, only employ the observing group. We find that the effect of the peer observation compared to small and large villages are significant. In villages with fewer households, peer observation seems especially pronounced (0.627, p-value=0.00). The magnitude and significance vanishes when we use only large villages - those having above 147 households (0.010, p-value=0.98). Results are available upon request. In addition, it seems that women in small villages tend to opt for the tasty treat more than in the larger villages. It is statistically significant at the 1% level. It seems that in small villages where people are more likely to know each other better and form relationships, peer effects are stronger than in large villages.

In the next step, we would like to investigate whether peer effects or the choice of the tasty treat is especially pronounced in the morning than in the afternoon. Taking the entire sample, we find that the group average still matters for the decision-making of the individual in terms of consumption. This is statistically significant at the 5% level. Hence, conducting the experiment in the morning or afternoon does not seem to matter for the full sample. It seems that peer effects on consumption decision are especially pronounced in the morning (0.74, p-value=0.01), however, in the afternoon they are less strong (0.35, p-value=0.11). Results are available upon request.

3.6 Conclusion

In a standard economic model of consumption choice, the effect of peers is largely ignored. Our study shows that peers observation has an effect on consumption of temptation goods.

We start by introducing a conceptual framework that introduces a cost if the individual makes a decision that deviates from that of the group. From this framework we can see that the demand function of the temptation good is less steep under peer observation. We can also derive a positive relationship between the average group choice and the individual choice.

In a lab-in-the-field experiment conducted in Thailand, we compare the decision between the tasty treat and an increasing amount of money. In the control group, respondents perform the experiment at the same time as their peers but without observing each other. In the treatment group, peers still make individual choices, but observe each other whilst doing it.

Due to the experimental nature and the large number of control variables, we can circumvent the identification problems normally associated in studies on peer effects. We find that standard deviations of those groups that observe each other are higher than for those groups that do not observe each other. At the same time, we show that individual choices are higher when the leave-out group mean is higher. Most importantly, we only observe this when the experiment is performed with peer observations. Hence, we provide clear evidence of peer effects and conclude that peer observation leads to conformity.

We further study the effect of familiarity with the product and find that peer observation can counteract the effect of a lack of knowledge of a product. Looking into treatment heterogeneities, we find that individuals with high cognitive ability, compared to their group, are less likely to choose the tasty treat, while the same effect is not to be found for low income, overconfident or high-income individuals.

We also test for the timing of the peer effects and find no significant changes in results.

Despite these findings, a lot of open questions remain that call for further research into peer effects and its effect of consumption choices. So far, there is no consensus on the "best" method for identifying peer effects, in part because models and methods must necessarily be case-specific. However, understanding the complexity of peer effects seems yet to be sufficiently explored. While initial estimates of such effects have been made, existing studies can and should be supplemented with spatial or non-linear analysis accounting for heterogeneous impacts. Manski (1993) provided a useful starting point for studies of peer effects, but many studies would benefit from thinking beyond a simple linear-in-means model. Furthermore, more research is needed that looks into the mechanisms behind peer effects. In more detail, a more structured experiment may be able to disentangle the effect of information and network externality and so explain why we find this conformity when peers observe each other. In addition, research could be done into the effect of key individuals within a group, that is to investigate who leads a group and who in a group follows but also how long peer effects may persist.

Table 1: Summary Statistics of Individual and Village Characteristics

Variable	Mean	Std. Dev.	Min.	Max.	N
Panel A: Individual Characteristics					
Female	0.597	0.491	0	1	543
Age	54.206	13.84	14	86	543
Married	0.826	0.379	0	1	541
Years of Schooling	5.629	3.105	1	17	529
Household Size	4.046	1.716	1	12	502
Number of children	1.129	1.058	0	7	513
Dependency Ratio	1.477	0.67	0	6	491
Farmer	0.685	0.465	0	1	502
Self-employed	0.056	0.23	0	1	502
Public Servant	0.02	0.14	0	1	502
Body Mass Index (BMI)	22.993	3.771	11.755	36.982	494
Per Capita Consumption	2397.427	1879.337	395.568	15638.178	548
Total Asset Value	10837.222	17783.051	-408.089	209066.234	502
Algebra Knowlegde	3.555	1.39	0	6	555
Overconfidence	0.365	0.482	0	1	298
Number of Animals	17.215	6.035	4	44	553
Cognitive Ability Measure	-0.028	1.404	-3.655	4.61	553
Panel B: Behavioral Variables					
Distance to district capital	15.964	9.676	2	40	550
Distance to Ubon	59.438	35.492	2	145	550
Number of village shocks	1.449	0.626	1	3	265
Number of households in village	167.007	89.453	55	813	535

Household size is the headcount of persons living in the household for at least 180 days. Body Mass Index is computed weight/height^2. Numeracy is the score out of six math questions (Details can be found in Appendix B). Number of animals named is the number of animals that someone can name in 60 seconds. Overconfident is a dummy that is unity if the respondent is overconfident. Cognitive Ability Measure is a PCA generated by performing principal component analysis on the numeracy score and the number of animals named in 60 seconds. Distance to Ubon is the average distance of the village to the provincial capital.

Table 2: Comparing Individual-Level Treatment and Control Group

Variable	**Control Group**	**Treatment Group**	**T-Test p-value**	**Wilcoxon Rank p-value**
Panel A: Individual Characteristics				
Female	0.57	0.62	0.28	0.27
Age	54.17	54.11	0.96	0.89
Married	0.80	0.85	0.13	0.13
Years of Schooling	5.61	5.67	0.85	0.56
Household Size	4.08	4.01	0.64	0.84
Number of children	1.22	1.01	0.02	0.06
Dependency Ratio	1.53	1.41	0.06	0.34
Farmer	0.69	0.69	0.98	0.98
Self-employed	0.05	0.06	0.60	0.59
Public Servant	0.03	0.01	0.28	0.28
Body Mass Index (BMI)	23.03	22.93	0.77	0.75
Per Capita Consumption	2299.92	2507.79	0.20	0.48
Total Asset Value	10699.97	11095.22	0.81	0.20
Algebra Knowledge	3.55	3.57	0.85	0.58
Number of Animals	17.22	17.20	0.97	0.94
Overconfidence	0.38	0.43	0.24	0.53
Cognitive Ability Measure	-0.03	0.04	0.97	0.95
Panel B: Behavioral Variables				
Distance to district capital	16.16	15.67	0.55	0.76
Distance to Ubon	65.05	53.68	0.00	0.00
Number of village shocks	1.47	1.41	0.48	0.76
Number of households in village	163.23	171.78	0.27	0.82
N (Individuals)	552			

The Table reports t-test and Wilcoxon ran sum test between treatment and control groups. Household size is the headcount of persons living in the household for at least 180 days. Body Mass Index is computed weight/height. Numeracy is the score out of six math questions (Details can be found in Appendix B). Number of animals named is the number of animals that someone can name in 60 seconds. Overconfident is a dummy that is unity if the respondent is overconfident. Cognitive ability, pca, is the score generated by performing principal component analysis on the numeracy score and the number of animals named in 60 seconds. Distance to district town/provincial capital is the average distance of the village to the provincial capital in kilometers.

Table 3: Comparing Observing and Non-Observing Peer Groups

Group Mean	**Observing Groups**	**Non-Observing Groups**	**T-Test p-value**	**Wilcoxon Rank p-value**
Female	0.58	0.63	0.30	0.24
Age	54.32	54.18	0.91	0.50
Married	0.81	0.85	0.14	0.12
Years of Schooling	5.73	5.68	0.69	0.86
Household size	4.09	4.02	0.65	0.64
Number of Children	1.23	1.02	0.03	0.03
BMI	23.09	23.07	0.95	0.82
Consumption	7.55	7.62	0.26	0.22
Feeling	2.22	2.26	0.64	0.52
Overconfident	0.43	0.43	0.95	0.98
Cognitive ability	-0.03	-0.02	0.94	0.76
N (Groups)	126			

The Table reports T-Test and Wilcoxon Rank test between observing and non-observing peer groups. Control Variables stay the same with the exception of feeling which asks how the respondent feels today before the start of the experiment. It is coded from 1(very good) to 5 (very bad).

Table 4: Comparing Outcomes for Observing and Non-Observing Peer Groups

Outcome PayTT	**Observing Groups**	**Non-Observing Groups**	**T-Test p-value**	**Wilcoxon Rank p-value**
Mean	2.94	2.93	0.91	0.70
Standard Deviation	2.26	1.70	0.00	0.00
Group maximum	5.74	4.93	0.01	0.04
Group minimum	0.68	1.21	0.03	0.11
N (Groups)	126			

The Table reports difference of the outcome choice between observing and non-observing peer groups. We use the payTT which is the last row subjects choose the tasty treat before swichting to money as the outcome variable. Group minimum is the lowest switching point within the group. Group maximum is the highest switching point within a group.

Table 5: Group-Level Treatment Effect on PayTT

	(1) **Mean PayTT**	(2) **Std.Dev. PayTT**	(3) **Max PayTT**	(4) **Min PayTT**
Peer Treatment	0.042	-0.477***	-0.111*	0.635**
	(0.14)	(-2.77)	(-1.67)	(2.26)
Mean Female	-1.346**	-0.695**	-0.408***	-0.480
	(-2.40)	(-2.02)	(-2.94)	(-0.94)
Mean Consumption	0.744	-0.293	0.056	1.193**
	(1.37)	(-0.96)	(0.50)	(2.11)
Mean Age	-0.013	-0.016	-0.003	0.004
	(-0.73)	(-1.33)	(-0.84)	(0.17)
Mean Cognitive Ability	-0.106	-0.119	-0.055	0.090
	(-0.49)	(-1.06)	(-1.17)	(0.49)
Mean Married	-0.825	-0.738	-0.324**	-0.368
	(-1.06)	(-1.60)	(-1.97)	(-0.47)
Mean No. of children	0.036	0.033	0.057	0.260
	(0.12)	(0.20)	(0.77)	(0.97)
Mean Schooling	0.072	0.084	0.032	-0.077
	(0.80)	(1.38)	(1.62)	(-0.86)
Mean Household Size	0.020	-0.057	-0.000	0.033
	(0.11)	(-0.60)	(-0.00)	(0.20)
Mean Feeling	-0.201	-0.030	-0.012	-0.132
	(-0.72)	(-0.17)	(-0.16)	(-0.56)
Mean Overconfidence	0.396	-0.185	-0.062	0.715
	(0.61)	(-0.52)	(-0.45)	(1.15)
Mean BMI	-0.015	0.045	-0.011	-0.057
	(-0.22)	(0.99)	(-0.69)	(-1.10)
Constant	-0.409	5.170**	2.03**	-7.826
	(-0.09)	(2.13)	(2.29)	(-1.53)
R-Squared	0.08	0.18		
Observations	126	126	126	126

The Table reports regression results with clustered standard errors in parenthesis. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. Peer observation is a dummy that is 1 if the game is played with peers observing each other. Mean (Variables) are the average group composition in the observing groups. Column 1 and 2 report OLS estimates. Columns 3 and 4 show poisson results.

Table 6, Panel A: Group Average Without Self and Individual choice on PayTT

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Last row TT	**Last row TT**	**Last row TT**	**Last row TT**	**Last row TT**	**Last row TT**	**Last row TT**
	All	All	Peer Treatment	Peer Treatment	Single Treatment	Single Treatment	All
Group Mean without Self	0.444^{***}	0.403^{***}	0.670^{***}	0.627^{***}	0.008	0.031	0.304^{*}
	(0.08)	(0.10)	(0.07)	(0.13)	(0.18)	(0.20)	(0.12)
Peer*Group Mean Without Self							0.162^{*}
							(0.08)
Group Average Characteristics (exluding the individual)	No	Yes	No	Yes	No	Yes	Yes
Individual Characteristics	No	Yes	No	Yes	No	Yes	Yes
Constant	1.59^{***}	2.88	0.91^{***}	2.47	2.84^{***}	7.05	3.41
	(0.25)	(2.25)	(0.21)	(3.79)	(0.53)	(4.40)	(2.34)
R-Squared	0.08	0.12	0.25	0.30	0.00	0.12	0.14
Observations	537	442	256	203	278	236	439

The Table reports OLS regression results with clustered standard errors in parenthesis. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. Group Average controls include all the controls from Table 3 but excluding the individual.

Table 6, Panel B: Group Average Without Self and Individual choice on MaxTT

	(1) **Last row TT** All	(2) **Last row TT** All	(3) **Last row TT** Peer Treatment	(4) **Last row TT** Peer Treatment	(5) **Last row TT** Single Treatment	(6) **Last row TT** Single Treatment	(7) **Last row TT** All
Group Mean without Self	0.156**	0.186**	0.261***	0.468***	-0.006	0.072	0.141*
	(0.05)	(0.06)	(0.06)	(0.12)	(0.09)	(0.09)	(0.07)
Peer*Group Mean Without Self							0.077
							(0.05)
Group Average Characteristics (exluding the individual)	No	Yes	No	Yes	No	Yes	Yes
Individual Characteristics	No	Yes	No	Yes	No	Yes	Yes
Constant	-1.71***	0.66	-2.09***	5.95	-1.24***	0.29	0.92
	(0.17)	(1.71)	(0.25)	(3.45)	(0.26)	(2.73)	(1.77)
Pseudo R-Squared	0.03	0.13	0.11	0.36	0.00	0.17	0.13
Observations	537	442	256	203	278	236	439

The Table reports Probit regression results with clustered standard errors in parenthesis. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. Group Average controls include all the controls from Table 3 but excluding the individual.

Table 6, Panel C: Group Average Without Self and Individual choice on NoTT

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Last row TT	**Last row TT**	**Last row TT**	**Last row TT**	**Last row TT**	**Last row TT**	**Last row TT**
	All	All	Peer Treatment	Peer Treatment	Single Treatment	Single Treatment	All
Group Mean without Self	-0.147**	-0.153**	-0.247***	-0.276**	0.023	-0.006	-0.115
	(0.05)	(0.06)	(0.07)	(0.09)	(0.08)	(0.09)	(0.07)
Peer*Group Mean Without Self							-0.078
							(0.05)
Group Average Characteristics (exluding the individual)	No	Yes	No	Yes	No	Yes	Yes
Individual Characteristics	No	Yes	No	Yes	No	Yes	Yes
Constant	-0.28	-0.32	-0.02	0.87	-0.79***	-2.53	-0.36
	(0.15)	(1.32)	(0.20)	(1.78)	(0.24)	(2.81)	(1.32)
Pseudo R-Squared	0.02	0.07	0.08	0.20	0.00	0.11	0.07
Observations	537	442	256	203	278	236	439

The Table reports Probit regression results with clustered standard errors in parenthesis. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. Group Average controls include all the controls from Table 3 but excluding the individual.

Table 7: Familiarity with the Tasty Treat

	(1) **Last row TT** All	(2) **Last row TT** All	(3) **Last row TT** Peer Treatment	(4) **Last row TT** Peer Treatment	(5) **Last row TT** Single Treatment	(6) **Last row TT** Single Treatment
Group Mean without Self	0.442**	0.331*	0.827***	0.841***	-0.384*	-0.482*
	(0.15)	(0.17)	(0.13)	(0.17)	(0.22)	(0.26)
Unfamiliarity with TT	-0.125	-0.452	0.917	1.05	-2.101***	-2.591***
	(0.58)	(0.66)	(0.58)	(0.86)	(0.74)	(0.92)
Peer*Unfamiliarity	-0.008	0.081	-0.203	-0.164	0.479**	0.55**
	(0.16)	(0.18)	(0.18)	(0.25)	(0.22)	(0.29)
Group Average Characteristics (exluding the individual)	No	Yes	No	Yes	No	Yes
Individual Characteristics	No	Yes	No	Yes	No	Yes
R-Squared	0.08	0.13	0.26	0.31	0.02	0.15
Observations	537	442	256	203	278	235

The Table reports OLS regression results with clustered standard errors in parenthesis. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. Group Mean controls include all the controls from Table 3.

Table 8: The Effect of Cognitive Ability on the Demand of TT

	(1) **Last row TT** Peer Treatment	(2) **Last row TT** Peer Treatment	(3) **Last row TT** Peer Treatment	(4) **Last row TT** Peer Treatment	(5) **Last row TT** Peer Treatment	(6) **Last row TT** Peer Treatment
Group Mean without Self	0.681***	0.782***	0.692***	0.670***	0.670***	0.653***
	(0.06)	(0.08)	(0.14)	(0.07)	(0.07)	(0.12)
High Aptitude		-0.791*	0.382	-0.251		
		(0.38)	(0.58)	(0.77)		
Peer*High Aptitude			-0.383*	-0.200		
			(0.16)	(0.19)		
Low Aptitude				-0.241	-0.301	-0.062
				(0.39)	(0.61)	(0.95)
Peer*Low Aptitude					0.021	-0.161
					(0.18)	(0.31)
Group Average Characteristics (exluding the individual)	No	No	Yes	No	No	Yes
Individual Characteristics	No	No	Yes	No	No	Yes
R-Squared	0.27	0.29	0.32	0.26	0.26	0.31
Observations	256	256	197	256	256	197

The Table reports OLS regression results with clustered standard errors in parenthesis. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. Group Mean controls include all the controls from Table 3.

Table 9: Comparing Observing and Non-Observing Peer Groups (Std. Dev.)

Group Mean	**Observing Groups**	**Non-Observing Groups**	**T-Test p-value**	**Wilcoxon Rank p-value**
Female	0.42	0.42	0.87	0.85
Age	12.79	12.23	0.59	0.72
Married	0.31	0.25	0.18	0.22
Years of Schooling	2.24	2.22	0.95	0.95
Household Size	1.51	1.52	0.94	0.42
Number of children	1.23	1.02	0.03	0.14
Body Mass Index	3.32	3.51	0.54	
Per Capita Consumption	0.54	0.58	0.37	0.72
Feeling	0.79	0.77	0.71	0.72
Overconfidence	0.46	0.42	0.32	0.15
Cognitive Ability Measure	1.28	1.19	0.29	0.36
N (Groups)	126			

The Table reports T-Test between observing and non-observing peer groups.

Table 10: Timing of Experiment and Peer Effects

	(1) **Last row TT** Peer Treatment	(2) **Last row TT** Single Treatment	(3) **Last row TT** All	(4) **Last row TT** All	(5) **Last row TT** All
Group Mean without Self	0.612***	-0.021	0.310*	0.332*	0.212
	(0.14)	(0.20)	(0.12)	(0.13)	(0.16)
Morning Dummy	-0.291	0.001	-0.192	-0.664	-0.682
	(0.35)	(0.42)	(0.24)	(0.59)	(0.59)
Peer*Group Mean Without Self			0.151		0.161*
			(0.08)		(0.08)
Morning*Group Mean Without Self				0.132	0.174
				(0.20)	(0.19)
Group Average Characteristics (exluding the individual)	Yes	Yes	Yes	Yes	Yes
Individual Characteristics	Yes	Yes	Yes	Yes	Yes
R-Squared	0.31	0.12	0.14	0.13	0.14
Observations	197	235	432	435	432

The Table reports OLS regression results with clustered standard errors in parenthesis. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. Group Mean controls include all the controls from Table 3.

Table 11: Sessions of Experiment and Peer Effects

	(1)	(2)	(3)	(4)	(5)
	Last row TT	**Last row TT**	**Last row TT**	**Last row TT**	**Last row TT**
	All	All	Peer Treatment	Peer Treatment	All
Group Mean without Self	0.441***	0.402***	0.674***	0.628***	0.033
	(0.08)	(0.10)	(0.07)	(0.13)	(0.19)
Session 1 Dummy	0.061	0.041	1.175	-1.333*	-1.743**
	(0.20)	(0.23)	(1.14)	(0.64)	(0.63)
Session 2 Dummy	-0.024	0.000	1.132	-1.204	-1.762**
	(0.22)	(0.29)	(1.09)	(0.72)	(0.60)
Peer*Group Mean Without Self					0.623*
					(0.21)
Group Average Characteristics (exluding the individual)	No	Yes	No	Yes	Yes
Individual Characteristics	No	Yes	No	Yes	Yes
R-Squared	0.08	0.12	0.26	0.31	0.17
Observations	537	435	256	197	432

The Table reports OLS regression results with clustered standard errors in parenthesis. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. Group Mean controls include all the controls from Table 3.

A Appendix
Experimental Instructions

We would now like to play a game with you in which you have to choose between some tasty goods or money. At the end of the game you can keep either the tasty goods or the money. We will ask you to choose between the two options 7 times. Each time we ask you, we increase the amount of money. The amount of tasty goods will always be the same. The enumerator will write down your choice each time we ask you. After the game, we will draw a number from a bag. This determines which of the two options you get. The tasty good will be given to you straight after the game. The money, however, will be given to you at the end of the whole survey. You will only receive one option. Either money or tasty good.

Example: No.3 is drawn from the bag. For the third decision you chose the tasty treat, so you will get the tasty treat immediately. *Enumerator put tasty good on the table.*

Enumerator will present the tasty good and ask the following question. **Please estimate the price of the tasty treat in the market.**

Price of tasty treat ____________ (THB)

Enumerator tells respondent that the price of the tasty present is THB 40 and put up the sign that shows the price.

Please choose!

Row	Tasty Good	Tick Box	Money
1	Tasty Good		10 THB
2	Tasty Good		20 THB
3	Tasty Good		30 THB
4	Tasty Good		40 THB
5	Tasty Good		50 THB
6	Tasty Good		60 THB
7	Tasty Good		70 THB

What is the maximum you would to pay for the tasty good? ____________ (THB)

Now chance will decide! Please draw a number. Number drawn: ____________ (THB)

B Appendix
Measurement of Numeracy and Overconfidence

	Details
Questions	Description
Word fluency	I would like you to name as many different animals as you can in 60 seconds.
Numeracy Q.1	What is 45 + 72?
Numeracy Q.2	You have 4 friends and you want to give each friend sweets. How many sweets do you need?
Numeracy Q.3	What is 5% of 200?
Numeracy Q.4	You want to buy a bag of rice that costs 270 Baht, You only have one 1000 Baht note. How much change will you get?
Numeracy Q.5	In a sale, a shop is selling all items at half price. Before the sale a mattress costs 3000 Baht. How much will the mattress cost in the sale?
Numeracy Q.6	A second-hand motorbike dealer is selling a motorbike for 12000 Baht. His is two thirds of what it costs new. How much did the motorbike cost new?
Overconfidence	How many of the 6 math's questions above, do you think you have answered correctly?

Bibliography

Abdellaoui, M., Baillon, A., Placido, L., and Wakker, P. (2011). The rich domain of uncertainty: Source functions and their experimental implementation. *American Economic Review*, 101(2):695–723.

Agarwal, S. and Mazumder, B. (2013). Cognitive abilities and household financial decision making. *American Economic Journal: Applied Economics*, 5(1):193–207.

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

Andersen, S., Harrison, G. W., Lau, M. I., and Rutström, E. E. (2008b). Lost in state space: Are preferences stable? *International Economic Review*, 49(3):1091–1112.

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

Anderson, L. R. and Mellor, J. (2008). Predicting health behaviors with an experimental measure of risk preference. *Journal of Health Economics*, 27(5):1260–1274.

Andersson, O., Tyran, J. R., Wengström, E., and Holm, H. J. (2013). Risk aversion relates to cognitive ability: Fact or fiction? *IFN Working Paper No. 964, Lund University.*

Angrist, J. (2014). The perils of peer effects. *Labour Economics*, 30(1):98–108.

Angrist, J. and Lang, K. (2004). Does school integration generate peer effects? Evidence from Boston's metco program. *American Economic Review*, 94(5):1613–1634.

Arrow, K. J. (1974). The use of unbounded utility functions in expected-utility maximisation: Response. *Quarterly Journal of Economics*, 88(1):136–138.

Baecker, A. and Mechtel, M. (2014). Peer effects in cheating on task performance. *Unpublished Manuscript.*

Bandiera, O., Barankay, I., and Rasul, I. (2010). Social incentives in the workplace. *Review of Economic Studies*, 77(2):417–458.

Bandiera, O. and Rasul, I. (2006). Social networks and technology adoption in northern Mozambique. *Economic Journal*, 116(514):869–902.

Banerjee, A. (1992). A simple model of herd behavior. *Quarterly Journal of Economics*, 107(3):797–817.

Banerjee, A. and Mullainathan, S. (2010). The shape of temptation: Implications for the economic lives of the poor. *Working Paper No. 257.*

Banerjee, A. V. and Duflo, E. (2007). The economic lives of the poor. *Journal of Economic Perspectives*, 21(1):141–168.

Barsky, R. B., Juster, T. F., Kimball, M. S., and Shapiro, M. D. (1997). Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *Quarterly Journal of Economics*, 112(2):537–579.

Bchir, M. A. and Willinger, M. (2013). Does the exposure to natural hazards affect risk and time preferences? Some insights from a field experiment in Peru. *Lameta Working Paper, University of Montpellier.*

Beauchamp, J., Cesarini, D., and Johannesson, M. (2012). The psychometric properties of measures of economic risk preferences. *Working Paper, Harvard University.*

Becketti, S., Gould, W., Lillard, L., and Welch, F. (1988). The panel study of income dynamics after fourteen years: An evaluation. *Journal of Labor Economics*, 6(4):472–492.

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

Berg, J., Dickhaut, J., McCabe, K., and Smith, V. (2005). Risk preference instability across institutions: A dilemma. *Proceedings of the National Academy of Sciences*, 102(11):4209–4214.

Bernheim, B. D. (1994). A theory of conformity. *Journal of Political Economy*, 102(5):841.

Bertrand, M. and Morse, A. (2011). Information disclosure, cognitive biases, and payday borrowing. *Journal of Finance*, 66(6):1865–1893.

Bertrand, M. and Morse, A. (2013). Trickle-down consumption. *NBER Working Paper Series 18883*.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1998). Learning from the behavior of others: Conformity, fads, and informational cascades. *Journal of Economic Perspectives*, 12(3):151–170.

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

Bolton, G. E. and Ockenfels, A. (2000). Erc: A theory of equity, reciprocity, and competition. *American Economic Review*, 90(1):166–193.

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

Brunnermeier, M. K. and Nagel, S. (2008). Do wealth fluctuations generate time-varying risk aversion? Micro-evidence on individuals' asset allocation. *American Economic Review*, 98(3):713–736.

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

Burns, W. J., Peters, E., and Slovic, P. (2012). Risk perception and the economic crisis: A longitudinal study of the trajectory of perceived risk. *Risk analysis : An official publication of the Society for Risk Analysis*, 32(4):659–677.

Bursztyn, L., Ederer, F., Ferman, B., and Yuchtman, N. (2014). Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions. *Econometrica*, 82(4):1273–1301.

Cai, H., Chen, Y., and Fang, H. (2009). Observational learning: Evidence from a randomized natural field experiment. *American Economic Review*, 99(3):864–882.

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

Calvet, L. E. and Sodini, P. (2014). Twin picks: Disentangling the determinants of risk-taking in household portfolios. *Journal of Finance*, 69(2):867–906.

Camerer, C. F. and Hogarth, R. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, 19(1):7–42.

Cameron, C. and Miller, D. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources, forthcoming.*

Cameron, L. and Shah, M. (2015). Risk-taking behavior in the wake of natural disasters. *Journal of Human Resources*, 50(2):484–515.

Carrell, S. E., Hoekstra, M., and West, J. E. (2011). Is poor fitness contagious? *Journal of Public Economics*, 95(7-8):657–663.

Charness, G. and Viceisza, A. (2012). Comprehension and risk elicitation in the field: Evidence from rural Senegal. *Working Paper, UC Santa Barbara.*

Charness, G. and Villeval, M. C. (2009). Cooperation and competition in intergenerational experiments in the field and the laboratory. *American Economic Review*, 99(3):956–978.

Chiappori, P. A. and Paiella, M. (2011). Relative risk aversion is constant: Evidence from panel data. *Journal of the European Economic Association*, 9(6):1021–1052.

Choi, S., Fisman, R., Gale, D., and Kariv, S. (2007). Consistency and heterogeneity of individual behavior under uncertainty. *American Economic Review*, 97(5):1921–1938.

Chuang, Y. and Schechter, L. (2015). Stability of experimental and survey measures of risk, time, and social preferences: A review and some new results. *Journal of Development Economics*, 117(7):151–170.

Clotfelter, C. T. and Cook, P. J. (1990). On the economics of state lotteries. *Journal of Economic Perspectives*, 4(4):105–119.

Cohn, A., Engelmann, J., Fehr, E., and Maréchal, M. A. (2015). Evidence for countercyclical risk aversion: An experiment with financial professionals. *American Economic Review*, 105(2):860–885.

Cole, S., Sampson, T., and Zia, B. (2011). Prices or knowledge? What drives demand for financial services in emerging markets? *Journal of Finance*, 66(6):1933–1967.

Conley, T. G. and Udry, C. R. (2010). Learning about a new technology: Pineapple in Ghana. *American Economic Review*, 100(1):35–69.

Crosetto, P. and Filippin, A. (2013). A theoretical and experimental appraisal of five elicitation methods. *SOEP Paper No.548, DIW Berlin.*

Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 47(2):448–474.

Cutler, D. and Glaeser, E. (2005). What explains differences in smoking, drinking, and other health related behaviors? *American Economic Review*, 95(2):238–242.

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

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

Deck, C., Lee, J., Reyes, J. A., and Rosen, C. C. (2009). A failed attempt to explain within subject variation in risk taking behavior using domain specific risk attitudes. *Journal of Economic Behavior and Organization*, 87(3):1–124.

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

Dercon, S. and Krishnan, P. (2000). In sickness and in health: Risk sharing within households in rural Ethiopia. *Journal of Political Economy*, 108(4):688–727.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review*, 100(3):1238–1260.

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

Donkers, B., Melenberg, B., and van Soest, A. (2001). Estimating risk attitudes using lotteries: A large sample approach. *Journal of Risk and Uncertainty*, 22(2):165–195.

Duflo, E., Dupas, P., and Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review*, 101(5):1739–1774.

Duflo, E. and Saez, E. (2003). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *Quarterly Journal of Economics*, 118(3):815–842.

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

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

Falk, A., Dohmen, T., Huffman, D., Sunde, U., and Becker, A. (2014). An experimentally validated preference survey module. *mimeo*.

Falk, A. and Ichino, A. (2006). Clean evidence on peer effects. *Journal of Labor Economics*, 24(1):39–57.

Fehr, E. and Schmidt, K. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114(3):817–868.

Fitzgerald, J., Gottschalk, P., and Moffitt, R. (1998). An analysis of sample attrition in panel data: The Michigan panel study of income dynamics. *Journal of Human Resources*, 33(2):251–299.

Foster, A. D. and Rosenzweig, M. R. (1995). Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of Political Economy*, 103(6):1176–1209.

Gardner, M. and Steinberg, L. (2005). Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: An experimental study. *Developmental psychology*, 41(4):625–635.

Gaudecker, H.-M. v., van Soest, A., and Wengstroem, E. (2011). Heterogeneity in risky choice behavior in a broad population. *American Economic Review*, 101(2):664–694.

Gertler, P. and Gruber, J. (2002). Insuring consumption against illness. *American Economic Review*, 92(1):51–70.

Giorgi, G. d., Pellizzari, M., and Redaelli, S. (2010). Identification of social interactions through partially overlapping peer groups. *American Economic Journal: Applied Economics*, 2(2):241–275.

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

Godlonton, S. and Thornton, R. (2012). Peer effects in learning HIV results. *Journal of development economics*, 97(1):118–129.

Goeree, J. K. and Yariv, L. (2011). An experimental study of collective deliberation. *Econometrica*, 79(3):893–921.

Gollier, C. and Pratt, J. W. (1996). Risk vulnerability and the tempering effect of background risk. *Econometrica*, 64(5):1109–1123.

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

Gunter, B. and Furnham, A. (1998). *Children as Consumers: A Psychological Analysis of the Young People's Market*. Routledge, London.

Guryan, J., Kroft, K., and Notowidigdo, M. J. (2009). Peer effects in the workplace: Evidence from random groupings in professional golf tournaments. *American economic journal: Applied economics*, 1(4):34–68.

Gusio, L., Sapienza, P., and Zingales, L. (2014). Time-varying risk aversion. *National Bureau of Economic Research, Working Paper No.19284*.

Hardeweg, B., Menkhoff, L., and Waibel, H. (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 Camerer, C. F. (1994). The predictive utility of generalized expected utility theories. *Econometrica*, 62(6):1251–1289.

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

Harrison, G. W., Lau, M. I., and Rutström, E. E. (2007). Estimating risk attitudes in Denmark: A field experiment. *Scandinavian Journal of Economics*, 109(2):341–368.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–161.

Hershey, J. C. and Schoemaker, P. J. H. (1985). Probability versus certainty equivalence methods in utility measurement: Are they equivalent? *Management Science*, 31(10):1213–1231.

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

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

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

Hirshleifer, D. and Shumway, T. (2003). Good day sunshine: Stock returns and the weather. *Journal of Finance*, 58(3):1009–1032.

Holm, H. J., Opper, S., and Nee, V. (2013). Entrepreneurs under uncertainty: An economic experiment in China. *Management Science*, 59(7):1671–1687.

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

Imberman, S. A., Kugler, A. D., and Sacerdote, B. I. (2012). Katrina's children: Evidence on the structure of peer effects from hurricane evacuees. *American Economic Review*, 102(5):2048–2082.

Isaac, R. M. and James, D. (2000). Just who are you calling risk averse? *Journal of Risk and Uncertainty*, 20(2):177–187.

Jalan, J. and Ravallion, M. (1999). Are the poor well-insured? Evidence on vulnerability to income risk in rural China. *Journal of Development Economics*, 58(1):61–81.

Jappelli, T. and Pistaferri, L. (2010). The consumption response to income changes. *Annual Review of Economics*, 2:479–506.

Kahnemann, D. and Amos, T. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292.

Kimball, M. S., Sahm, C. R., and Shapiro, M. D. (2008). Imputing risk tolerance from survey responses. *Journal of the American Statistical Association*, 103(483):1028–1038.

Knight, F. (1921). *Risk, Uncertainty, and Profit.* Riverside Press, Cambridge.

Kremer, M. and Levy, D. (2008). Peer effects and alcohol use among college students. *Journal of Economic Perspectives*, 22(3):189–206.

Kremer, M. and Miguel, E. (2007). The illusion of sustainability. *The Quarterly Journal of Economics*, 122(3):1007–1065.

Levin, I. P., Schneider, S. L., and Gaeth, G. J. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes*, 76(2):149–188.

Levitt, S. D. and List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives*, 21(2):153–174.

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

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

Loomes, G. and Sugden, R. (1998). Testing different stochastic specifications of risky choice. *Economica*, 65(260):581–598.

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

Malmendier, U. and Nagel, S. (2011). Depression babies: Do macroeconomic experiences affect risk taking? *Quarterly Journal of Economics*, 126(1):373–416.

Mangyo, E. and Park, A. (2011). Relative deprivation and health: Which reference groups matter? *Journal of Human Resources*, 46(3):459–481.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *Review of Economic Studies*, 60(3):531.

Manski, C. F. (2000). Economic analysis of social interactions. *Journal of Economic Perspectives*, 14(3):115–136.

Mas, A. and Moretti, E. (2009). Peers at work. *American Economic Review*, 99(1):112–145.

Meier, S. and Sprenger, C. D. (2015). Stability of time preferences. *Review of Economics and Statistics*, 97(2):273–286.

Miguel, E. and Kremer, M. (2004). Worms: Identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, 72(1):159–217.

Moffit, R. (2001). *Policy Interventions, Low-Level Equilibria, and Social Interactions*. Brookings Institution Press, Washington DC.

Moskowitz, T. J. and Vissing-Jörgensen, A. (2002). The returns to entrepreneurial investment: A private equity premium puzzle? *American Economic Review*, 92(4):745–778.

Mossin, J. (1968). Aspects of rational insurance purchasing. *Journal of Political Economy*, 76(4):553–568.

Oster, E. and Thornton, R. (2012). Determinants of technology adoption: Peer effects in menstrual cup take-up. *Journal of the European Economic Association*, 10(6):1263–1293.

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

Quiggin, J. (2003). Background risk in generalized expected utility theory. *Economic Theory*, 22(3):607–611.

Rabe-Hesketh, S. and Skrondal, A. (2012). *Multilevel and Longitudinal Modeling Using Stata (Third Edition).* Stata Press.

Reynaud, A. and Couture, S. (2012). Stability of risk preference measures: Results from a field experiment on french farmers. *Theory and Decision*, 73(2):203–221.

Sacerdote, B. (2001). Peer effects with random assignment: Results for dartmouth roommates. *Quarterly Journal of Economics*, 116(2):681–704.

Sahm, C. (2012). How much does risk tolerance change? *Quarterly Journal of Finance*, 2(4):1–53.

Slovic, P. (1964). Assessment of risk taking behavior. *Psychological Bulletin*, 61(3):220–233.

Steinberg, L. and Cauffman, E. (1996). Maturity of judgment in adolescence: Psychosocial factors in adolescent decision making. *Law and Human Behavior*, 20(3):249–272.

Stigler, G. J. and Becker, G. S. (1977). De gustibus non est disputandum. *American Economic Review*, 67(2):76–99.

Sutter, M., Kocher, M. G., Glätzle-Rützler, D., and Trautmann, S. T. (2013). Impatience and uncertainty: Experimental decisions predict adolescents' field behavior. *American Economic Review*, 103(1):510–531.

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

Townsend, R. M. (1994). Risk and insurance in village India. *Econometrica*, 62(3):539.

Tversky, A. and Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481):453–458.

Veblen, T. (1899). *The Theory of the Leisure Class: an Economic Study of Institutions*. Macmillan Publishers, London.

Vella, F. and Verbeek, M. (1999). Two-step estimation of panel data models with censored endogenous variables and selection bias. *Journal of Econometrics*, 90(239-263).

Vieider, F. M., Lefebvre, M., Bouchouicha, R., Chmura, T., Hakimov, R., Krawczyk, M., and Martinsson, P. (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.

Vieider, F. M., Nghi, T., and Martinsson, P. (2013). Risk preferences and development revisited: A field experiment in Vietnam. *WZB Discussion Paper, No. 401*.

Viscusi, W. K. and Hersch, J. (2001). Cigarette smokers as job risk takers. *Review of Economics and Statistics*, 83(2):269–280.

Viscusi, W. K., Phillips, O. R., and Kroll, S. (2011). Risky investment decisions: How are individuals influenced by their groups? *Journal of Risk and Uncertainty*, 43(2):81–106.

Volk, S., Thöni, C., and Ruigrok, W. (2012). Temporal stability and psychological foundations of cooperation preferences. *Journal of Economic Behavior and Organization*, 81(2):664–676.

Voors, M. J., Nillesen, E. E. M., Verwimp, P., Bulte, E. H., Lensink, R., and Soest, D. P. V. (2012). Violent conflict and behavior: A field experiment in Burundi. *American Economic Review*, 102(2):941–964.

Wakker, P. P. (2010). *Prospect Theory: For Risk and Ambiguity.* Cambridge University Press.

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

Weber, M., Weber, E. U., and Nosic, A. (2013). Who takes risks when and why: Determinants of changes in investor risk taking. *Review of Finance*, 17(3):847–883.

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