Appendices

To accompany “Estimating Risky Behavior with Multiple-Item Risk Measures”

Appendix A Information about Risk Elicitation Methods

A.1 Implementation of Risk Eliciation Tasks

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.

Implementation of the risk elicitation methods contained two village visits per day, one in the morning and one in the afternoon. The possibility that information had spread between villages seems unlikely because most of the distance between villages (18km on average). 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. 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 CSL 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 et al. (2007). 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, the largest of the three provinces in Northeastern Thailand.

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 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. Hypothetical investment question: 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 Option

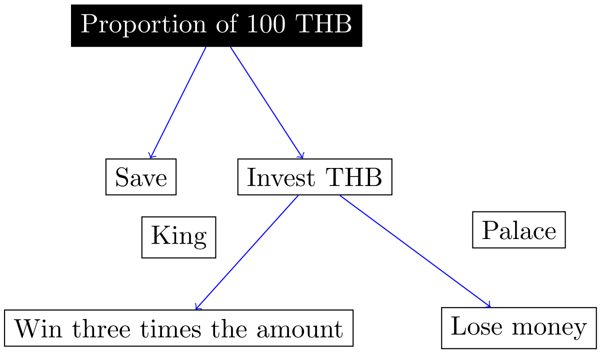
	Choice	Coin	Stakes	Tick Option
Option	1	King	50	
		Palace	50	
Option	2	King	90	
		Palace	30	
Option	3	King	130	
		Palace	10	
Option	4	King	170	
		Palace	-10	
Option	5	King	210	
		Palace	-30	

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 Option

	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.

<u>**Individual Characteristics**</u>

Female 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 takes the value 1 if married and 0 otherwise.

Household size are individuals living in the household for at least 180 days.

Log per capita income is the natural logarithm of household income per day divided by OECD adult equivalents AE (AE = 1+0.7*(adults-1) + 0.5*children).

<u>**Behavioral Variables**</u>

Farmer takes the value 1 for being a farmer and 0 otherwise.

Self-employed takes the value 1 for being self-employed and 0 otherwise.

Lottery expenditure is the total amount of household expenses for lotteries in the last 12 months.

Future lottery expenditure denotes the question: How much money will you spend for the next drawing of the state lottery?

Investment is 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 takes the value of 1 if the household borrowed since 2011 and 0 otherwise.

Borrowing for business takes the value 1 if respondent applied for a credit for business purposes and 0 otherwise.

Risk-mitigating measures 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.

Health insurance takes the value of 1 if the household holds additional health insurance next to the public one and 0 otherwise.

Accident insurance the value of 1 if the household holds an additional accident insurance and 0 otherwise.

Body mass index is computed by weight/height^2.

Overweight 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.

Cognitive aptitude

We asked five basic algebra questions and tested 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 income. Age has an ambiguous relationship with risk attitude. It seems that younger people are more risk seeking in the SF as well as the SI. This is in line with the results found in (Dohmen et al., 2011). This result is different in both CS 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. Furthermore, we find that higher income is positively significant in explaining risk tolerance which is statistically significant in the SI task. Overall, significant relations have the expected signs; with the exceptions of age.

Table C.1: Determinants of Risk Measures

	(1) SG	(2) SF	(3) SI	(4) CE	(5) CS	(6) CSL	(7) CI
Female	0.377	-0.008	-3.959	-1.229	-0.102	-0.019	-4.221
	(0.29)	(0.29)	(2.50)	(0.68)	(0.17)	(0.15)	(2.88)
Age	0.010	-0.027**	-0.355***	-0.021	0.013**	0.016***	-0.175
	(0.01)	(0.01)	(0.07)	(0.03)	(0.01)	(0.01)	(0.13)
Height	0.010	-0.010	0.162	-0.079	-0.006	0.006	-0.224
	(0.02)	(0.02)	(0.19)	(0.04)	(0.01)	(0.01)	(0.17)
Years of schooling	0.062	0.023	2.558	0.136	0.035	-0.029	0.513
	(0.04)	(0.05)	(2.53)	(0.11)	(0.02)	(0.02)	(0.40)
Marital status	0.815**	0.932**	6.320***	0.694	0.099	-0.040	3.187
	(0.36)	(0.39)	(2.28)	(0.78)	(0.16)	(0.15)	(3.36)
Household size	0.003	0.107	3.110	0.092	-0.027	0.021	0.030
	(0.07)	(0.07)	(4.31)	(0.17)	(0.03)	(0.03)	(0.69)
LPCI	-0.039	-0.005	1.893**	0.006	0.040	0.105	-0.402
	(0.13)	(0.12)	(0.82)	(0.30)	(0.07)	(0.06)	(1.27)
Constant	3.779	8.235**	24.361	20.594**	3.022	0.467	81.491***
	(3.44)	(3.41)	(21.43)	(7.95)	(2.14)	(1.60)	(30.58)
Observations	715	715	715	715	715	715	715
R-Squared	0.01	0.03	0.09	0.01	0.01	0.03	0.01
Estimator	OLS	OLS	OLS	OLS	OLS	OLS	OLS

Notes:
The dependent variables are risk elicitation methods. Columns (1-7) report estimates of least squares 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). LGPI is the log per capita income. It refers to the natural logarithm of household income 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% and 5% levels.

Appendix D Factor Analysis

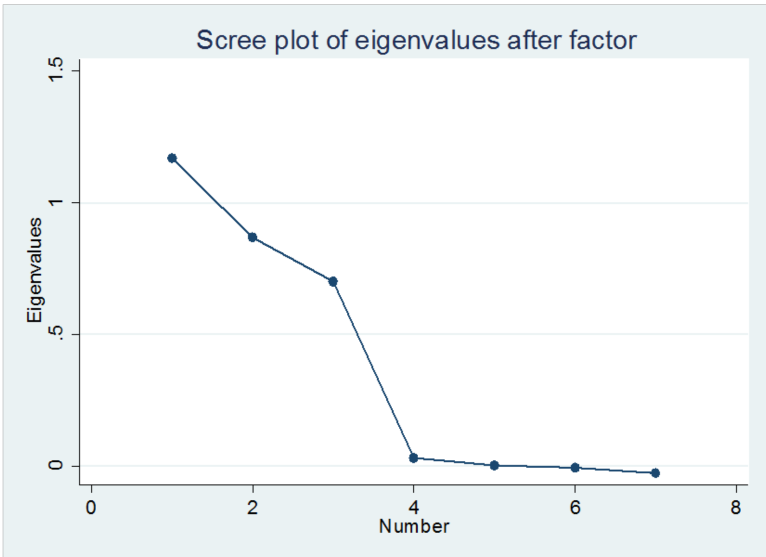
In the correlation matrix we saw varying degrees of correlation across risk elicitation methods. We aim now to reduce the number of variables and to detect structure in the relationships between risk items, employing a standard factor analysis. The Kaiser-Meyer-Olkin (KMO) measure verifies the sampling 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 SI. The second factor explains 32% of all variance and is dominated by the two choice set tasks. 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 SG and to some extent by SF 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 D.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 SI, Factor 2 is defined by the two choice set tasks, while Factor 3 is defined by the two survey questions.

Table D.1: Factor Analysis

Variable	Factor 1	Factor 2	Factor 3
SG			0.7548
SF			0.4852
SI	0.9591		
CE			
CSL		0.6216	
CS		0.7085	
CI			

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 D.1: Screeplot



D.1 Factor Analysis and Single-Item Risk Measures

Next, we consider each factor and its power in explaining the same behavioral items as above (Table D.2). 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 4 out of 11 cases. It is, for instance, able to explain risk-mitigating activities and the demand for health insurance policies at the 1% significance level. Factor 2, which captures the two CS items, is able to explain risky behavior in another area than Factor 1. Factor 3, relying on the two survey items, is unable to predict any behavior. Using factor analysis, we find that we can reduce our seven risk items to three factors.

Table D.2: Factor Analysis and Multiple-Item Risk Measures

	(1) Lottery (Expend/Inc)	(2) Future Lottery Expenditure	(3) Self-employment	(4) Farming	(5) Investment (Expend/Inc)	(6) Plan to Invest	(7) Borrowing	(8) Risk Mitigating	(9) Number of Insurance	(10) Health Insurance	(11) BMI
Factor 1			0.019**					-0.056***		-0.031***	0.387**
			(0.01)					(0.02)		(0.01)	(0.14)
Factor 2		29.320**						-0.062**	-0.202**		
		(12.44)						(0.03)	(0.08)		
Factor 3											
Observations	709	709	713	713	713	713	713	713	713	713	709
Average of 3 Items	0.030**	16.954***						-0.081***		-0.019**	0.351***
	(0.01)	(6.04)						(0.02)		(0.01)	(0.13)
Average of 2 Items		28.08**						-0.073***	-0.118**	-0.021**	0.316**
		(12.36)						(0.02)	(0.06)	(0.01)	(0.13)
Observations	711	711	715	715	715	713	715	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/Inc in column (1) is the total amount of household expenses on buying lottery tickets in the last 12 months as a share of total income. Expend/Inc in column (5) is the total amount of household expenses on investment 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 in years, and log per capita income. We employ the subsample with age of 17-79. Clustered errors on the village level are in parenthesis. *** and ** denote significance at the 1% and 5% levels.

Appendix E Robustness Tests

Robustness examinations cover five issues: First, we use further extensions of the behavioral items, including variations of consumption and income definitions (Section E.1). Motivation and results of three other robustness exercises are described in Sections E.2., E.3, and E.4. Finally, we report estimates on effect sizes of coefficients of the single-item and multiple-item risk measures in E.5.

E.1 Extension of Risky Behavior

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.

(1) We use the share of past lottery expenditure as part of consumption instead of income. (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. 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 E.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 15 out of 56 cases a single-item risk measure is able to explain behavior. Hence, the main picture drawn from Table 4 has not changed. Taking the total amount of investment, we still find that the CI task is significant. This also holds when we take the share of investment made by the household as part of total consumption. The CI task is statistically significant for all of the investment variables. Hence, variations of the investment variable shows that the previously significant risk items remain constant. For the indicator variable of borrowing in general, we found that the SF was successful in revealing borrowing behavior. This time, however, we find a significant and positive relationship with the SI and CI task at the 1% level in addition to the CSL. 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 15 out of 56 cases.

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

E.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 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. We test the assumption whether 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 E.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 CI and SI. The largest improvement can be seen between the CS experiments.

Essentially, from the results above we can infer that education slightly improves correlation between the experimental measures where probabilities are part of the task. Hence, understanding seems to play a role, yet it is not the decisive factor explaining low correlations.

E.3 Further Analysis of WTR (Gen)

The responses of our sample population to the SG 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 SG 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 manage to repeat the experiment with the same subjects in August. Ultimately, we have 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 SG item from spring 2013, we find that it is unable to explain any behavioral items. Using the average of all seven risk items with the SG item of the household survey; it is still significantly explains 7 out of 11 items. 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 SG from spring and summer 2013 to see whether results hold in Table 4. Table E.4, Panel B reports results. We find that the average of both SG items from the household panel survey and risk survey does not change results. The average is still able to explain the share of lottery expenditure at the 5% level.

E.4 Restricting the Sample to the Household Head

Table E.5 replicates the results of Table 4, however, using this time only the head of households as respondents. Using household heads reduces the observation to around 400 subjects. Out of these 400 subjects, 303 are male respondents. Household heads seem to be less risk tolerant compared to the spouse who often is a woman. The difference in risk attitude between men and women, 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.

Overall, our single-risk items are able to explain 14 out of 77 risky behavioral items. We do also find significant heterogeneity in terms of the strength of various risk items. While the SI and CI were able to explain four and three behavioral items in Table 4, other single-item risk measure in Table E.5 like the SG and CSL are now able to explain up to three items. Hence, it is more equally balanced. Except with the SG in column (7), all significant results have the expected signs.

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 only using male respondents. 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 performs better as it explains 3 behaviors. 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.

E.5 Effect Size

In order to ease the interpretation of the coefficients of the single-item risk measures in Table 4, we standardize variables following Long and Freese (2014). Table E.6 produces the standardized coefficients' output of the risk measures on risky behavior. **bStdX** presents standardized values of the explanatory variable on the outcome variable.

In column (1), we find that a one standard deviation increase in the normalized SG question produces, on average, a 0.021 increase in the % share of income spending on gambling whereas a one standard deviation increase in CS results in an increase of 0.025, holding other variables constant. In the second column, we follow the same procedure. We find that a one standard deviation increase in the CSL risk item (std.dev. 0.371) increases, on average, the future spending on lottery by 24.113 which has a higher impact compared to the CI and CS item. In column (3) a one standard deviation increase in CI results, on average, in an increase in the probability of being self-employed by 0.194. This effect is slightly reduced using the SI (0.130) SG (0.180). In column (4), we find a similar magnitude of the CE item on the probability of being a farmer compared to the SG item. Concerning financial behavior, column (5- 7) we find that the a standard deviation increase in the risk items CI, CE and SF increases the probability for risky financial behavior by 0.190, 0.149, and 0.237, respectively. In contrast, in column (8) we find that a one standard deviation increase in the SI item (0.216) from the mean (0.503), decreases on average, the probability of planning substantial investment in the future by 0.135 which is similar to the CSL item (-0.129). In the last column, we find that a one standard deviation increase in the CI risk measure increases the variation of the BMI measure by 0.365, holding other items constant. However, one needs to be careful in interpretation and comparing magnitudes of the risk items because risky behavioral variables differ by construction and therefore cannot be directly compared.

As for the multiple-item risk measures with six or seven items, we find that standardized coefficients are comparable to the beta values in Table 4, as values for the latter are standardized. For the other outcome variables where we use a probit model, we reported marginal effects in Table 4. Now, we report the values of the one standard deviation and find that values are similar.

Table E.1: Extensions of Risky Behavior

	(1) Lottery (Expend/Cons)	(2) Log Investment	(3) Total Investment	(4) Investment (Expend/Cons)	(5) Borrowing (Business)	(6) Number of Risk-Mitigating	(7) Accident Insurance	(8) Over-Weight
SG	0.011**				-0.027**			
	(0.00)				(0.01)			
SF								
SI		0.010***			0.001***			0.000***
		(0.00)			(0.00)			(0.00)
CE								
CSL	0.016**				0.013**		-0.013**	0.025**
	(0.00)				(0.01)		(0.01)	(0.01)
CS							-0.016***	
							(0.01)	
CI		0.010***	35.508**	1.226***	0.001***	-0.003**		
		(0.00)	(15.56)	(0.46)	(0.00)	(0.00)		
Observations	711	390	715	715	715	715	715	715
Average of 7 Items	0.026***	0.194***			0.037***	-0.156***		
	(0.01)	(0.07)			(0.01)	(0.04)		
Observations	709	389	713	713	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 column (1) is the total amount of household expenses on buying lottery tickets in the last 12 months as a share of total consumption. Expend/Cons in column (4) is the total amount of household expenses on investment in the last 12 months as a share of total consumption. See Appendix B for details. Controls include Female, Age, Height, Marital status, Household size, education in years and log per capita income. We employ the subsample with age of 17-79. Clustered errors on the village level are in parenthesis. *** and ** denote significance at the 1% and 5% levels.

Table E.2: Factor Analysis and Multiple-Risk Item Measures (Robustness)

	(1) Lottery (Expend/Cons)	(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		0.227***			0.030***			0.045**
		(0.07)			(0.01)			(0.02)
Factor 2							-0.034***	
							(0.01)	
Factor 3	0.035							
	(0.02)							
Observations	709	389	713	713	713	713	713	713
Average of 3 Items	0.047**	0.148***			0.030***	-0.130**		
	(0.02)	(0.07)			(0.01)	(0.04)		
Average of 2 Items		0.176***			0.026***	-0.027**	-0.227***	0.044**
		(0.07)			(0.01)	(0.05)	(0.01)	(0.02)
Observations	711	390	713	713	715	715	715	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 column (1) is the total amount of household expenses on buying lottery tickets in the last 12 months as a share of total consumption. Expend/Cons in column (4) is the total amount of household expenses on investment in the last 12 months as a share of total consumption. See Appendix B for details. Controls include Female, Age, Height, Marital status, Household size, education in years and log per capita income. We employ the subsample with age of 17-79. Clustered errors on the village level are in parenthesis. *** and ** denote significance at the 1% and 5% levels.

Table E.3: Spearman's Rank Correlations for Subsample

	SG	SF	SI	CE	CS	CSL	CI
SG	1.000						
SF	a) 0.381*** b) 0.401***	1.000					
SI	a) 0.128** b) 0.033	a) 0.041 b) 0.092	1.000				
CE	a) -0.005 b) 0.115	a) 0.070 b) 0.006	a)0.027 b) 0.054	1.000			
CSL	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		
CS	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	
CI	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% and 5% levels.

Table E.4: SG Variations and Risky Behavior

	(1) Lottery (Expend/Inc)	(2) Future Lottery Expenditure	(3) Self-employment	(4) Farming	(5) Investment (Expend/Inc)	(6) Plan to Invest	(7) Borrowing	(8) Risk Mitigating	(9) Number of Insurance	(10) Health Insurance	(11) BMI
					A: SG replacement						
SG											
Average of 7 Items	0.030***	10.013***			0.208***	0.062**	0.065**	-0.054**		-0.026**	
	(0.01)	(4.91)			(0.08)	(0.02)	(0.03)	(0.02)		(0.02)	
Observations	477	481	482	485	485	485	485	485	485	485	482
					B: SG over time						
SG (HH survey)	0.039**										
	(0.02)										
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/inc in column (1) is the total amount of household expenses on buying lottery tickets in the last 12 months as a share of total income. Expend/Inc in column (5) is the total amount of household expenses on investment 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 in years and log per capita income. We employ the subsample of 17-79. Clustered errors on the village level are in parenthesis. *** and ** denote significance at the 1% and 5% levels.

Table E.5: Household Head and Risky Behavior

	(1) Lottery (Expend/Inc)	(2) Future Lottery Expenditure	(3) Self-employment	(4) Farming	(5) Investment (Expend/Inc)	(6) Plan to Invest	(7) Borrowing	(8) Risk Mitigating	(9) Number of Insurance	(10) Health Insurance	(11) BMI
SG	0.008**		0.012**				-0.014**				
	(0.00)		(0.01)				(0.01)				
SF							0.020**	-0.017**			
							(0.01)	(0.01)			
SI										0.001**	
										(0.00)	
CE				-0.011**							
				(0.00)							
CSL	0.019**								-0.140**	-0.020***	
	(0.01)								(0.05)	(0.01)	
CS		13.052***						-0.044***			
		(5.65)						(0.02)			
GP					0.011***			-0.002**			
					(0.00)			(0.00)			
Observations	392	392	396	394	396	396	396	396	396	396	396
Average of 7 Items			0.032**					-0.090***		-0.037**	
			(0.02)					(0.03)		(0.01)	
Observations	391	391	395	395	395	395	395	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/Inc in column (1) is the total amount of household expenses on buying lottery tickets in the last 12 months as a share of total income. Expend/Inc in column (5) is the total amount of household expenses on investment 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 in years, and log per capita income. We employ the subsample with age of 17-79. Clustered errors on the village level are in parenthesis. *** and ** denote significance at the 1% and 5% levels.

Table E.6: Effect Size Estimates

	(1) Lottery (Expend/Inc) bstdX	(2) Future Lottery Expenditure bstdX	(3) Self-employment bstdX	(4) Farming bstdX	(5) Investment (Expend/Inc) bstdX	(6) Plan to Invest bstdX	(7) Borrowing bstdX	(8) Risk Mitigating bstdX	(9) Number of Insurance bstdX	(10) Health Insurance bstdX	(11) BMI bstdX
SG	0.021		0.180								
SF							0.237			-0.147	
SI			0.131					-0.135		-0.281	0.365
CE				-0.189		0.149					
CSL		24.113						-0.129			
CS	0.025	10.004							-0.188		
GP		16.293	0.194		0.190						
Observations	392	392	396	394	396	396	396	396	396	396	396
Average of 7 Items	0.026	12.751				0.104	0.127	-0.1554		-0.1975	
Average of 6 Items	0.054	13.952	0.196			0.111		-0.1471		-0.1871	
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/Inc in column (1) is the total amount of household expenses on buying lottery tickets in the last 12 months as a share of total income. Expend/Inc in column (5) is the total amount of household expenses on investment 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 in years, and log per capita income. We employ the subsample with age of 17-79. Clustered errors on the village level are in parenthesis. *** and ** denote significance at the 1% and 5% levels.