Norwegian University of Life Sciences (NMBU)

Can the risky investment game predict real world investments?

Stein T. Holden and Mesfin Tilahun

Norwegian University of Life Sciences
Centre for Land Tenure Studies

Centre for Land Tenure Studies Working Paper 05/22
ISBN: 978-82-7490-302-9

Highlights

Can the risky investment game predict real world investments?

Stein T. Holden,Mesfin Tilahun

- We use a field experiment to assess whether investment in the risky investment game predicts real world investments
- Correlations with four types of real world investments are assessed in a within-subject design
- Investment in the risky investment game is not significantly correlated with real world investments
- The field experiment reveals that investment in the game is associates with substantial measurement error
- Cognitive memory of the game played one year earlier is associated with higher investment in the game

Can the risky investment game predict real world investments?[⋆,⋆⋆]

Stein T. Holden[a], Mesfin Tilahun[a,b]

[a]School of Economics and Business, Norwegian University of Life Sciences, Towerbuilding, Ås, 1432, Viken, Norway
[b]Department of Economics, Mekelle University, Mekelle, , Tigray, Ethiopia

ARTICLE INFO

Keywords:
risky investment game
field experiment
prediction
measurement error
cognitive memory
Ethiopia
JEL codes
C93
D90

ABSTRACT

The incentivized risky investment game has become a popular tool in lab-in-the-field experiments for its simplicity and ease of comprehension compared to some of the more complex Multiple Choice List approaches that have been more commonly used in laboratory experiments. We use a field experiment to test whether the game can predict real world investments by the same subjects based on the assumption that the game can provide a reliable measure of risk tolerance and that risk tolerance is an important predictor of investment behavior. The results show that the game cannot predict investment behavior in our sample. There are two reasons for this. First, we find substantial measurement error and low correlation when the game is repeated one year later for the same subjects. Measurement error is so large in our sample that the "obviously related instrumental variable" (ORIV) approach of Gillen, Snowberg and Yariv (2019) could not remedy the problem. Second, the game appears to suffer from low asset integration due to narrow bracketing, explaining its limited predictive power and the failure to detect attenuation bias due to measurement error. Subjects' cognitive memory of the game played one year earlier is strongly positively related to investment intensity in the game and this result is much enhanced when correcting for the endogeneity of cognitive memory.

1. Introduction

It is generally believed that risk preferences are crucial determinants of investment behavior but there are few studies that have investigated this (Beauchamp, Cesarini and Johannesson, 2017). Risk preferences are not directly observable and a variety of theories, tools and methods have been used to measure and estimate such preferences. Even though there have been many achievements in form of theoretical and methodological contributions to the literature on risk preferences, there is still no consensus about the appropriate choice of tools and their predictive power in real world settings (Gillen et al., 2019). Measurement error may explain weak predictive power and weak correlation between different risk preference measures (Gillen et al., 2019; Chuang and Schechter, 2015).

Our area of primary interest is the real behavior of poor people in developing countries and the importance of risk preferences for their investment behavior. Many tools used to measure risk preferences are complicated, may be hard to understand, and require substantial numeracy skills that may be beyond the capacity of subjects with limited education and numeracy skills. The use of such tools for "non-WEIRD" samples may result in large measurement errors and this could be one of the reasons for poor prediction power.

It has been suggested that simpler tools should be pursued and used for populations with less education (Dave, Eckel, Johnson and Rojas, 2010; Chuang and Schechter, 2015). Dave et al. (2010) found that subjects with weaker numeracy skills had more inconsistent and noisy responses with the Holt and Laury (2002) (HL) Multiple Choice List[1] (MCL) approach than the simpler ordered lottery game of Binswanger/Eckel/Grossman (BEG) (Binswanger, 1981; Eckel and Grossman, 2002) and that the latter was preferable for subjects with limited math skills. Their study used a fairly large sample of mostly young adults in Canada with on average 12 years of education. They also concluded that the simpler task may generate risk preference estimates that are more stable across time. Charness and Viceisza (2016) compared the simple one-shot version of the risky investment game (Gneezy, Leonard and List, 2009) game

[⋆]This paper is the result of two research projects funded by the Norwegian Agency for Development Cooperation (NORAD) and the Research Council of Norway.

*Corresponding author
stein.holden@nmbu.no (S.T. Holden)
www.steinholden.com (S.T. Holden)
ORCID(s): 0000-0001-7502-2392 (S.T. Holden); 0000-0003-3885-2017 (M. Tilahun)

[1]Often called Multiple Price Lists but we think Choice List is a better name as each row in a list typically presents two prospects and may not present an explicit price.

with the more complicated HL approach in rural Senegal, using seeds rather than money. They classify risk elicitation approaches in complex versus simple and incentivized versus non-incentivized, and the risky investment game and the BEG approach fall under the category of simple and incentivized games that may be more suitable for field experiments with non-students with limited education in developing countries. The literature review by Charness and Viceisza (2016) of the performance of the HL approach in the field in developing countries revealed much higher inconsistency rates than has been found with the same approach in developed countries. Their sample was small (45+46 subjects for the HL and the risky investment game) so they drew only tentative conclusions and see a need for more research to scrutinize the the different tools. However, they concluded that the HL task was not well understood. Dasgupta, Mani, Sharma and Singhal (2019) used a large student sample from India to compare the BEG approach and the risky investment game in a within-subject design. They found a good correlation across the two games in their student sample and also found that decisions in both games were correlated with willingness to compete in a third game. Using a CRRA utility function they found higher average risk aversion in the risky investment game than in the ordered lottery game[2]. Dasgupta et al. (2019) concluded that both the risky investment game and the ordered lottery game can be used to elicit behavior under uncertainty in laboratory and field environments.

Kimball, Sahm and Shapiro (2008) assessed the extent of response errors in hypothetical risk tolerance questions and found that the variability due to response errors greatly exceeded that from the true variation in risk tolerance in a test-retest in a sub-sample of their data. The limited predictive power of incentivized as well as hypothetical experimental measures of risk preferences may be due to such response errors. Gillen et al. (2019) assessed the extent of measurement error in preferences based on three different tools and showed that measurement errors were substantial and caused attenuation bias and demonstrated a method to remove or reduce measurement errors and the biases they may cause. We build on this in our assessment of the one-shot risky investment game in a developing country setting with a similarly sized sample of young members of rural business groups.

In this study we test the predictive power of the one-shot version of the risky investment game that was first used by Gneezy et al. (2009)[3]. The advantages of this game are that it is easy to comprehend, it is incentivized, it does not take long time to implement, and it can quite easily be incorporated in larger household surveys. Two limitations of the risky investment game are that a) it cannot, as a stand alone device, separate the effects of utility curvature, probability weighting and loss aversion and b) it cannot in the standard one-shot version remove noise (measurement error) from the investment decisions in the game. However, its simplicity may reduce the risk of there being much measurement error in the game.

Most studies that have assessed the correlations between risk preferences and investments have collected risk preference data after the investments were made (Chuang and Schechter, 2015) and rely on the assumption that risk preferences are stable over time. One example is Liu (2013) who used a risk preference experiment based on Prospect Theory (PT) (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992) to study how the risk preference parameters were correlated with the adoption process for a genetically modified pest-resistant Bt cotton variety during the previous decade. Holden and Quiggin (2017) assessed how risk preferences jointly with a climatic shock (drought) affected the adoption of drought-tolerant and other maize varieties using a rural sample of smallholder farmers in Malawi. Both these studies used more complex tools that derived Prospect Theory parameters from rural populations and found significant prediction power for the estimated parameters. Our question is whether simpler tools such as the risky investment game that is integrated in a survey can predict real world investment behavior?

In our study we utilize a fairly large sample that combines incentivized experimental data from the one-shot risky investment game and survey-based investment data collected one year later from the same subjects. The combined 2016 and 2017 experimental and survey sample consists of 966 young business group members belonging to 116 business groups in northern Ethiopia. A business group is a part-time formal (primary cooperative) business supported by the local government to create an additional livelihood opportunity for resident resource-poor young adults. We study the individual risk tolerance as revealed in the game and how it correlates with investments made by these members the following year, dividing investments into durable consumer goods, livestock, productive assets and other business investments. To assess the stability and reliability of the responses in the risky investment game we repeated the game in an hypothetical version one year later (test-retest assessment), the cognitive memory of the choice and outcome in the real game one year after it was played, and as a basis for predicting risk tolerance and correcting for measurement error based on the ORIV approach proposed by Gillen et al. (2019). Our study also speaks to the literature on cognitive ability and risk preferences (Burks, Carpenter, Goette and Rustichini, 2009; Lilleholt, 2019). Higher risk tolerance has

[2]This finding is consistent with the finding of significant endowment effects in the risky investment game (Holden and Tilahun, 2021).
[3]Tt is derived from the more complex version developed by Gneezy and Potters (1997) to study myopic loss aversion.

been associated with higher cognitive ability and more rational decisions. Our study contributes by assessing how a specific type of cognitive ability; the ability to memorize a game played one year earlier, including the choice made and the outcome. We also contribute by dealing with the potential endogeneity of cognitive ability and show that its correlation with risk tolerance is substantially higher after controlling for such endogeneity.

Our study contributes to the broader literature on the development of simple and reliable field experimental tools to elicit measures of risk tolerance and that can explain and predict behavior of resource-poor populations with limited education. It is of high policy relevance to develop and test such tools that in combination with survey data can help predict real world behavior. To our knowledge this is the first study to test the investment predictive power of the risky investment game by collecting investment data one year after the game was played and then revisit the game, the cognitive memory of the decision and lottery outcome of the game. Our paper contributes to the assessment of noise in the game through a test-retest and the influence of cognitive memory on risk preferences and consistency of choices in the game.

2. Experimental and survey design

We use data from a survey and field experiments implemented in the period July-August 2016 for 1134 subjects belonging to 119 business groups in five districts in Tigray Region in northern Ethiopia. The experimental data have previously been used in one of the treatments studying endowment effects associated with different framing in the game (Holden and Tilahun, 2021). We combine these data with survey data collected at the same time of the year (July-August) one year later from 966 of the same subjects (116 groups). The key survey questions and experimental protocol (in English) are provided in Appendix 1. All questions were translated to the local language, *Tigrinya*, and asked by experimental enumerators who recorded the data using tablets programmed using the software CSPro.

2.1. The risky investment game

The standard one-shot risky investment game (Gneezy et al., 2009) was used for all subjects without any treatment variation. Each subject was allocated an initial endowment $X = 30$ ETB[4]. The subjects were free to invest nothing, some, or all of the endowment=x (in multiples of 5 ETB) in a 50-50 lottery with the researcher tripling the the amount invested. This implies that a lucky winner gets $30 + 2x$ with probability 0.5 and $30 - x$ with probability 0.5. A risk neutral person should invest the full amount. We measure the share invested such that $0 \leq x/30 \leq 1$. The basic issue we want to study is whether this investment share ($r = x/X$) can be a good indicator of risk tolerance and can explain or predict real world investment, either alone or conditional on other variables that are likely to influence investment.

2.2. Survey and lab-in-the field experimental design

We initially selected 120[5] youth business groups out of a total of 742 such groups identified in a census covering the five districts in 2015 (Holden and Tilahun, 2018). Groups that were allocated land on a more permanent basis for their business were included in the study[6].

Up to 12 business group members from each group were sampled. The experiment and survey interviews were carried out in local schools during the school holiday. The interview of 12 members from a business group were carried out simultaneously by locating an enumerator[7] and a group member in one corner each of three classrooms with the members facing the enumerator sitting in the corner and with a desk between them. This was done to prevent communication between the group members and spillover effects. Twelve trained enumerators were therefore interviewing only one subject per group, thereby preventing the confounding of group characteristics with enumerator performance. In addition, three supervisors were orchestrating the survey, experiments, and interviews to prevent communication between groups as well. All experiments and interviews in one location were completed in one day.

[4]ETB is Ethiopian Birr, the local currency. 30 ETB was approximately equal to a daily wage rate in the study locations at the time of the survey.

[5]Four groups refused to participate for religious reasons as they associated our experiments with gambling.

[6]306 of the groups in the census were temporary as they were allocated allocated temporary mineral rights to allow them to extract and sell minerals till they reached a certain capital limit, when their mineral extraction right would be terminated.

[7]The enumerators were carefully selected and trained in the local university, then interviewing and doing the experiments in pairs as enumerator and subject, before they were doing field experiments and interviews with out-of-sample groups under close supervision.

Table 1
Descriptive statistics

	Mean	Median	SD	Min	Max	N
2016 Investment game and socio-economic variables						
Riskshare (Investment share in risky investment game 2016)	0.44	0.33	0.25	0	1	966
Sex (Female dummy)	0.32	0.00	0.47	0	1	966
Age, years	29.29	27.00	9.76	15	74	965
Education, years	5.38	6.00	3.94	0	17	965
Parents' farm size, *tsimdi* (0.25ha units)	2.27	2.00	2.16	0	16	958
Livestock endowment (Tropical livestock units)	1.24	0.70	1.64	0	16	966
Durable assets, number	1.34	1.00	1.46	0	9	966
Log(income in 1000 ETB)	1.64	1.79	1.08	0	5	966
2017 Investment variables						
Amount invested in durable consumer goods ETB	929	300	1,973	0	20,600	966
Amount invested in livestock ETB	3,225	1,058	5,058	0	50,000	966
Amount invested in other productive assets ETB	1,121	839	1,666	0	30,630	966
Amount invested other business activities ETB	4,154	0	14,545	0	170,000	966
Total investment last year (2016-17)	9,429	4,210	17,286	0	186,760	966
Riskshare17h (Risk share in risky investment game 2017)	0.43	0.33	0.28	0	1	953
Cognitive memory index	2.23	2	1.10	0	4	941
Outcome of lottery, risky investment game 2016	0.58	1	0.49	0	1	965
Idiosyncratic shock 2016-17, dummy	0.17	0.00	0.38	0	1	966

3. Descriptive statistics

Table 1 presents the descriptive statistics for the experiment and basic socioeconomic data collected in 2016 and the investment variables, the hypothetical game, the cognitive memory index for the real game played one year earlier, and an idiosyncratic shock variable collected in 2017, one year after the initial survey and experiment. Riskshare is the share out of 30 ETB that the subjects invested in the risky investment game. To further illustrate the distribution of the investments in the real investment game played in 2016, see Fig. 1. Fig. 2 shows the distribution in the hypothetical game played one year later by the same subjects. Fig. 1 shows that only 10% invested the full amount and with the mode and median at 10 ETB (riskshare=0.33). Fig.2 shows a similar distribution in the hypothetical game one year later. This indicates that only a small share of the sample are risk neutral. Table 1 shows that only 32% of the subjects are female and that the median level of completed education is 6 years. The subjects come from land-poor households as the median farm size for the parent households is only 0.5 ha (2 *tsimdi* which is below the average farm size in the area based on recent land registry data (Holden and Tilahun, 2020). Livestock is an important asset and source of livelihood in this semiarid area. Climate shocks are common in the area but were not a problem in the period 2016-17. About 17% of the sample suffered some form of idiosyncratic shock in the 2016-17 period.

4. Theoretical foundation and empirical strategy

Building on Prospect Theory (PT) the investment level in the game may be shown to possibly depend on how subjects weight the probability, the curvature of their utility or value function as well as loss aversion. Subjects are up-front in the game provided a monetary endowment or budget for the game. Two additional fundamental unobservable issues can influence decisions in the game. First, is the cash endowment provided integrated with the background cash or income or wealth of the subjects? Second, does the up-front allocation cause the subjects to change their reference point for judging whether the invested money is perceived to be in the gains domain or in the loss domain? The literature on endowment effects indicates that subjects can quickly shift reference points and this may contribute to explain endowment effects for monetary endowments like also has been found in the risky investment game (Holden and Tilahun, 2021). This may imply that loss aversion is attributed to potential losses in the game that therefore are given more weight than potential gains. Note also that PT assumes that only changes in income (gains and losses) matter and this implies that the monetary endowment provided in the game is not integrated with background income or wealth when making decisions in the game, implying no asset integration or narrow bracketing. Even early studies

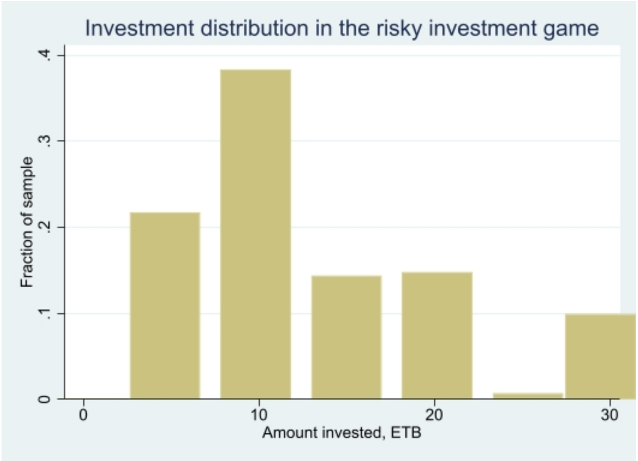

Figure 1: Investment distribution in the 2016 real risky investment game

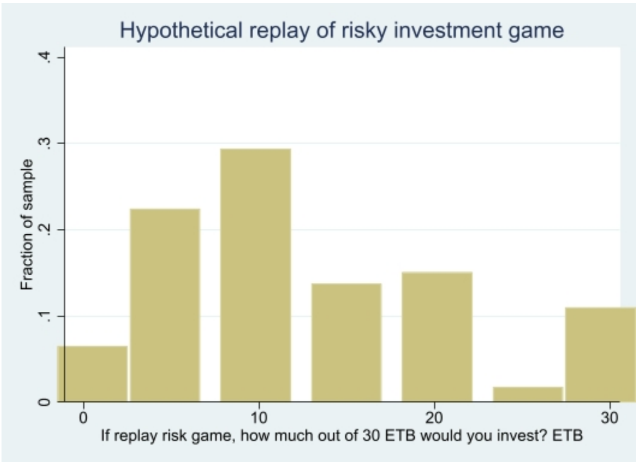

Figure 2: Investment distribution in the 2017 hypothetical risky investment game

of risk preferences found that respondents reacted much more to the states in the game than to variation in background wealth or income, pointing towards limited or no asset integration.

A model for maximizing behavior in the game that builds on PT and quickly changing reference points can therefore be expressed as follows (denoting loss aversion as λ) (Holden and Tilahun, 2021):

$$\max PT(T1) = w^{+}(0.5)v(2x) - w^{-}(0.5)\lambda v(|x|) \tag{1}$$

We may consider some special cases of this game. First, if the utility or value function is linear, respondents would invest all or nothing in the game. They would invest nothing if $\lambda \geq 2$ and the full amount otherwise. Second, optimism or pessimism may affect the probability weighting and cause the weighted probability of gain or loss to deviate from 0.5. Combined with a linear value or utility function this will still only affect whether subjects invest all or nothing. The fact that interior solutions are dominating in the game may imply that subjects have concave value functions in addition to being loss averse. Generally, risk tolerance in the game is assumed to increase with the degree of optimism in probability weighing, decrease with concavity of the value function and decrease with degree of loss aversion.

$$r = x/X = r(w^{+}(0.5), v(\cdot), \lambda) \tag{2}$$

So how is behavior in the game associated with the general investment behavior of subjects? First, if the underlying preference characteristics in form of loss aversion, utility/value function curvature and probability weighting are stable subject characteristics, behavior in the game reflects how a combination of these provides a measure of subjects' risk tolerance. If also real world investments are driven by the same stable preference characteristics, higher risk tolerance in the game should be positively correlated with real world investments. However, real world investments are not only driven by risk preferences but also by complementary resource endowments, skills, knowledge and budget constraints. We control for the business environment using business group fixed effects (g_g). The groups are spatially confounded to a specific location that characterizes the local environment in terms of group invariant resource, community, weather, institutional and market characteristics.

The different types of investments by each subject are not independent. It is likely that the different types of investments compete for limited resources while each may be positively related to the total investments of subjects. We assess this and the importance of the risk tolerance as measured with the risky investment game with a sequence of models starting with a parsimonious version that does not include the other investment types and total investment, only risk tolerance, group fixed effects and enumerator fixed effects (control for possible measurement error). We estimate a set of log-linear investment models for the four investment types and total investment by the subjects for investments made by our study subjects in the following year after risk tolerance was revealed with the game. After the parsimonious model, we estimate a model that includes additional controls. These are the other investment variables that are likely to be correlated, individual demographic and resource endowment variables:

$$I_{t,ij=k} = \alpha_0 + \alpha_1 r_{t-1,gi} + \alpha_2 I_{t,ij\neq k} + \alpha_3 z_{t-1,gi} + \alpha_4 E_d + g_g + \varepsilon_{gi} \tag{3}$$

where $I_{t,ij=k}$[8] represents investment type k, $r_{t-1,gi}$ represents the investment level in the risky investment game played in 2016, $z_{t-1,gi}$ represents individual socio-economic (demographic and resource endowment) characteristics recorded in 2016, E_d represents enumerator fixed effects, g_g represents business group fixed effects, and ε_{gi} is the error term.

One obvious limitation of the last set of models where other investment types and resource endowments are included as controls, is that these additional RHS variables are likely to be correlated with risk tolerance and this generates an endogeneity bias. We return to this issue of endogeneity bias in the ORIV models which are used to obtain consistent estimates of the effect of risk tolerance on investment below (equation 7).

As tests of asset integration in the risky investment game and a possible attrition effect[9], the investment in the game is regressed on the predetermined individual socio-economic characteristics that were collected at the same time as the game was played in 2016:

$$r_{t-1,gi} = r_0 + r_1 z_{t-1,gi} + r_2 E_d + g_g + \epsilon_{gi} \tag{4}$$

We have used a new (hypothetical) round of the risky investment game in the 2017 survey round to assess a) how closely investments in the real game in 2016 are related to the investments in the hypothetical game one year later, b) how this is influenced by the subjects' cognitive memory[10] (cm) of the game one year earlier. We estimate a model to assess this while imposing business group and enumerator fixed effects[11] as controls (equation 5).

$$r_{t,gi} = \tau_0 + \tau_1 r_{t-1,gi} + \tau_2 cm_{t,gi} + \tau_3 E_{t-1,d} + \tau_4 E_{t,d} + g_g + \epsilon_{gi} \tag{5}$$

Next, we assess the correlations between the stated investment in the hypothetical game and socioeconomic variables. This is a replication of the model in equation (4) except that the dependent variable is replaced with the new hypothetical response variable and the addition of the cognitive memory index and double enumerator fixed effects as additional RHS variables (equation (6)):

$$r_{t,gi} = r'_0 + r'_1 z_{t-1,gi} + r'_2 cm_{t,gi} + r'_3 E_{t-1,d} + r'_4 E_{t,d} + g_g + \epsilon_{gi} \tag{6}$$

[8]We denote 2016 as $t-1$ to emphasize the timing of the game and the variables used. The real investment variables were for the period 2016-2017 and are therefore flow variables denoted as t variables based on the time when they were collected. We suppress the timing subscript for the enumerator, business group variables, and the error term to keep the notation simple.

[9]The model was run on the full sample from 2016 as well as for the select sample for which investment data were collected one year later.

[10]We constructed a 5-level (0-4) index for how much of the game they remember and the correctness of their memory of the game.

[11]We include enumerator dummies for both years as controls for possible measurement errors in the risky investment games in both years.

We cannot rule out that cognitive memory in equation (5) is endogenous and correlated with $r_{t-1,gi}$ (lagged risk tolerance) based on the literature demonstrating such a connection (Burks et al., 2009; Lilleholt, 2019).

As our identification strategy we utilize the random outcome of the real investment game played in 2016 and propose that random winners (rw) (win/loss dummy) in the game remember it better than random losers. The outcome (win/loss dummy) is therefore used as an instrument to predict the cognitive memory ($cmhat$). Subjects are more likely to remember details about a game where they won money. At the same time, such a small stakes game[12] should not affect the risk preferences of the subjects, implying that we have an instrument that is valid and orthogonal on the choice in the hypothetical game. The instrument was also found to be strong. Equation (7) outlines the panel IV model.

$$\begin{aligned} r_{t,gi} &= \eta_0 + \eta_1 z_{t-1,gi} + \eta_2 I_{t,ij} + \eta_3 cmhat_{t,gi} + \eta_5 E_{t-1,d} + \eta_6 E_{t,d} + g_g + v_{gi}, \\ \text{instrumenting } cm_{t,gi} &: \\ cm_{t,gi} &= \gamma_0 + \gamma_1 rw_{t-1,gi} + \upsilon_{gi} \end{aligned} \tag{7}$$

Lastly, we utilize the two measures of investment levels in the risky investment game to assess and correct for measurement error based on the "obviously related instrumental variables" (ORIV) approach for the case where the explanatory variable risk tolerance is measured with error (Gillen et al., 2019). This is a stacked regression where the two measures of risk tolerance instrument for each other. This gives two observations per subject. The approach generates consistent estimates of γ in equation (8) where n is the number of subjects(Gillen et al., 2019). We get consistent standard errors by correcting them for clustering on subjects.

$$\begin{aligned} &\begin{pmatrix} I_{t,ij} \\ I_{t,ij} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \gamma \begin{pmatrix} r_{t-1,gi} \\ r_{t,gi} \end{pmatrix} + \xi_{gi}, \quad \text{instrumenting} \\ &\begin{pmatrix} r_{t-1,gi} \\ r_{t,gi} \end{pmatrix} \quad \text{with} \quad M = \begin{pmatrix} r_{t,gi} & 0_n \\ 0_n & r_{t-1,gi} \end{pmatrix} \end{aligned} \tag{8}$$

We estimate this model with 2SLS with enumerator and group fixed effects separately for the five investment variables. To take into account that the investment variables are not independent but correlated with each other, we estimate a stacked model for the four investment variables, leaving out the total investments. This is equivalent to having four imperfect measures of investment intensity per subject that jointly may be influenced by risk tolerance that is estimated with error. We therefore apply the model Gillen et al. (2019) proposed for estimation when there are errors in outcome as well as explanatory variables. This implies stacking four models like in equation (8) on top of each other and yielding eight observations per subject. The instrumentation matrix becomes $8n * 8$ in size. With this approach we get consistent estimates of the effect of risk tolerance. By clustering standard errors on subjects we also obtain consistent standard errors. Their efficiency depends on the degree of measurement error and thereby the strength of the instruments which is an empirical issue.

5. Results and discussion

The parsimonious panel investment models are presented in Table 2. The investment share in the risky investment game is insignificant but with a positive sign in all five models. This indicates that the risk tolerance revealed in this game has little influence on real world investment even after local business environment characteristics are controlled for with group fixed effects.

To further scrutinize this result we introduce additional controls in terms of other investment categories and individual demographic and resource endowment characteristics. The results are presented in Table 3. As expected, the results demonstrate that the different types of investments compete for household resources as demonstrated by their significant negative signs, while investment in each investment type is positively correlated with total investment. However, again the risk tolerance as measured with the risky investment game is insignificant in all models and even has a negative sign in two of the models. However, this may be due to endogeneity as risk tolerance is expected to be correlated with the RHS investment and resource endowment characteristics. Several of the individual socioeconomic controls were significantly correlated with investments and some significant heterogeneity across investment models is observable.

[12] Note that e.g. the average livestock investment by the subjects over the previous year was 100 times larger.

Table 2
Parsimonious loglinear panel investment models with group and enumerator fixed effects

VARIABLES	(1) logconsinv1617	(2) loganiminv1617	(3) logprodassinv1617	(4) logbusinv1617	(5) logsuminv1617
Riskshare 2016	0.495	0.048	0.045	0.516	0.295
	(0.483)	(0.520)	(0.397)	(0.638)	(0.288)
Group fixed effects	Yes	Yes	Yes	Yes	Yes
Enumerator fixed effects	Yes	Yes	Yes	Yes	Yes
Constant	4.019***	4.130***	4.731***	1.555***	7.209***
	(0.378)	(0.455)	(0.339)	(0.447)	(0.299)
Observations	966	966	966	966	966
R-squared	0.051	0.036	0.097	0.052	0.050
Number of business groups	116	116	116	116	116

Log(Consumer goods investment+1)=logconsinv1617, Log(Livestock investment+1)=loganiminv1617,
Log(Productive assets investment+1)=logprodassinv1617, Log(Business investment+1)=logbusinv1617,
Log(Total investment+1)=logsuminv1617
Cluster-robust standard errors in parentheses, clustering on business groups, *** p<0.01, ** p<0.05, * p<0.1.

As an additional test of asset integration in the risky investment game, we regressed the investment level in the game on the subjects' socioeconomic characteristics collected at the same time as the game was played in 2016. We did this for the full sample from 2016 as well as the select sample for which we obtained investment data one year later. Table 4 shows the results. Only the female dummy variable was significant and with a negative sign, in line with findings in other studies that have used this game(Charness and Viceisza, 2016; Dasgupta et al., 2019). The lack of significance of income and resource endowment variables may be an indication of narrow bracketing/no asset integration in this game. This is also in line with Prospect Theory. The provision of up-front cash in the game may contribute to such narrow bracketing on the game budget that is not integrated with background income or resource endowments. If loss aversion is the primary driver in the game, the implication is that real world investments by our sample population to a small extent are constrained by loss aversion. We recommend further work to assess how investments in the game are associated with the different prospect theory parameters but it is beyond the scope of this paper.

Can measurement error explain the results above? Even if the subjects in our study understand the game well, they may be whimsical in their responses. Next, as a test of the reliability of the game, we assessed the degree of correlation between the stated investment in the hypothetical risky investment game in 2017 with the investment level in the real game played one year earlier and how this is associated with their cognitive memory of the choice and outcome in the game played one year earlier. We see in Fig. 1 and 2 and in Table 1 that the distributions are quite close to each other at the aggregate level. However, this may not say much about the within-subject stability over time. To assess this we take the raw correlation coefficient for the investment levels in the game at the two points in time for the balanced sample of subjects. It is only 0.135, and this points towards substantial measurement error in the game. This is much lower than that found by Gillen et al. (2019) using the Caltech sample consisting of undergraduate students. The correlation increased from 0.03 for the subjects with cognitive memory score=1, to 0.20 for those with the highest memory score=4 (perfect memory of investment level and outcome in the game). This is substantially lower than Gillen et al. (2019) found even across different types of risk tolerance measures. They found that the risky investment game[13] was the one that was most highly correlated with the other risk tolerance elicitation measures. Our results indicate that many change their response even if they remember what they decided last time. The limited correlation with background variables point towards measurement error rather than instability of risk tolerance as the culprit.

We want to further examine the correlations between the two risk tolerance measures from the risky investment game, see Table 5. In the regression models we use enumerator fixed effects for both years to control for possible measurement error due to potential enumerator bias. In addition, we control for group fixed effects (location and group specific effects that may affect investment decisions). Model (1) in Table 5 assesses how the risk investment level in the 2016 real game and the cognitive memory of it contribute to explain the investment level in the 2017 hypothetical

[13] They call it the "Project measure".

Table 3
Investment models with extended list of individual controls

VARIABLES	(1) logconsinv1617	(2) loganiminv1617	(3) logprodassinv1617	(4) logbusinv1617	(5) logsuminv1617
Riskshare	0.225	-0.323	-0.113	0.266	0.059
	(0.450)	(0.444)	(0.334)	(0.541)	(0.181)
logconsinv1617		-0.162***	-0.106***	-0.080**	0.172***
		(0.037)	(0.027)	(0.040)	(0.020)
loganiminv1617	-0.164***		-0.122***	-0.118***	0.223***
	(0.036)		(0.024)	(0.044)	(0.013)
logprodassinv1617	-0.211***	-0.239***		-0.209***	0.289***
	(0.054)	(0.049)		(0.059)	(0.027)
logbusinv1617	-0.065**	-0.093**	-0.084***		0.160***
	(0.033)	(0.036)	(0.023)		(0.013)
logsuminv1617	0.914***	1.169***	0.770***	1.058***	
	(0.059)	(0.075)	(0.048)	(0.107)	
Sex (Female dummy)	-0.114	-0.534*	0.047	-0.008	-0.263*
	(0.297)	(0.273)	(0.170)	(0.280)	(0.143)
Age, years	-0.043**	-0.032*	0.008	0.008	0.008
	(0.017)	(0.018)	(0.012)	(0.023)	(0.006)
Education, years	0.121***	-0.007	-0.017	0.059	-0.010
	(0.033)	(0.034)	(0.023)	(0.042)	(0.013)
Farm size of parents	0.101*	0.072	0.074*	0.044	-0.041*
	(0.052)	(0.048)	(0.044)	(0.063)	(0.022)
Tropical livestock units	0.179**	-0.012	0.085	0.072	-0.046
	(0.078)	(0.092)	(0.062)	(0.125)	(0.037)
Durable assets, number	0.059	0.215*	0.167**	0.078	-0.009
	(0.106)	(0.110)	(0.069)	(0.136)	(0.046)
Log(individual income)	0.275**	0.327**	0.070	-0.011	0.002
	(0.133)	(0.136)	(0.091)	(0.165)	(0.061)
Group fixed effects	Yes	Yes	Yes	Yes	Yes
Enumerator fixed effects	Yes	Yes	Yes	Yes	Yes
Constant	-1.204	-2.151**	-0.496	-5.043***	4.002***
	(0.791)	(0.913)	(0.562)	(1.127)	(0.360)
Observations	957	957	957	957	957
R-squared	0.255	0.348	0.346	0.258	0.627
Number of business groups	116	116	116	116	116

Log(Consumer goods investment+1)=logconsinv1617, Log(Livestock investment+1)=loganiminv1617,
Log(Productive assets investment+1)=logprodassinv1617, Log(Business investment+1)=logbusinv1617,
Log(Total investment+1)=logsuminv1617
Cluster-robust standard errors in parentheses, clustering on business groups, *** p<0.01, ** p<0.05, * p<0.1.

game. The model shows that the correlation, after imposing these controls, is highly significant, but an investment share increase from 0 to 1 in the 2016 game only increases the investment share by 0.11 in the 2017 hypothetical game. The cognitive memory index is also highly significant and with a positive sign. One unit increase in the index raises the investment share by 0.025 which is a modest increase. The latter result is consistent with the finding that cognitive ability and risk tolerance are correlated (Burks et al., 2009; Lilleholt, 2019). Limited cognitive ability therefore can be one explanation for the low correlation between the two measures.

Model (2) in Table 5 is comparable to the models in Table 4, assessing the correlation between the hypothetical 2017 risk tolerance measure with the 2016 socioeconomic variables with the addition of the cognitive memory index. For the socioeconomic variables the gender dummy effect is almost identical to that for the real game played one year earlier

Table 4
Risky investment game: Test of asset integration

VARIABLES	(1) riskshare	(2) riskshare
Sex (Female dummy)	-0.052***	-0.065***
	(0.018)	(0.021)
Age, years	0.000	-0.000
	(0.001)	(0.001)
Education, years	0.003	0.002
	(0.003)	(0.003)
Farm size parents	0.003	0.004
	(0.004)	(0.004)
Livestock endowment	0.002	0.000
	(0.008)	(0.009)
Durable assets, number	0.005	0.009
	(0.009)	(0.010)
Log(Income), 1000ETB	-0.003	-0.010
	(0.008)	(0.009)
Group fixed effects	Yes	Yes
Enumerator fixed effects	Yes	Yes
Constant	0.357***	0.390***
	(0.050)	(0.059)
Observations	1,122	957
R-squared	0.089	0.091
Number of business groups	119	116

Cluster-robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

(model (2) in Table 4) while all endowment variables (except education) remain insignificant. Asset endowments, other than cognitive ability and education therefore seem to play no role in the decisions in the game.

Cognitive memory may, however, be endogenous and correlated with lagged risk tolerance based on findings in earler studies (Burks et al., 2009; Lilleholt, 2019). This, in combination with measurement error for the risk tolerance measures may explain an attenuation bias in the riskshare 2016 variable in model (1) in Table 5 and can also contribute to an endogeneity bias in the cognitive memory variable in models (1) and (2). With our identification strategy, utilizing the random outcome of the 2016 game as an instrument to predict the cognitive memory variable, model (3) shows a 10x increase in the coefficient for the cognitive memory variable[14]. This indicates a much stronger correlation between cognitive memory and risk tolerance than in the model that does not control for such bias. This is an important result and speaks to the literature on cognitive ability and risk preferences. Studies that have ignored this possible endogeneity may underestimate the effect of cognitive ability on risk tolerance.

The fact that the coefficient on the riskshare 2016 variable did not change much from model (1) to model (3) may indicate that model (3) did not adequately address the measurement error problem. Even the gender variable has become insignificant, possibly due to its correlation with the riskshare 2016 variable as well. Furthermore, all the real investment variables remained insignificant.

But we may ask one more question. Can the change in response in the game from 2016 to 2017 be due to a preference change? A recent and rapidly expanding literature has investigated whether risk preferences are stable over time or respond to shocks. The findings in this literature are mixed but many studies have detected preference changes after shocks. We included a shock variable in our study in form of a dummy variable for those that had been exposed to idiosyncratic shocks over the period between our two visits. 17% of our sample reported such shocks. This variable was added in our model (3) in Table 6. It was insignificant. Based on economic theory and this result we retain the

[14] A test for endogeneity revealed a robust score (chi2(1))=17.2, p=0.0000. A test for the strength of the instrument (win in the 2016 game) gave an F-value=16.1, demonstrating its strength.

Table 5
The 2017 risky investment (hypothetical) game and prediction

VARIABLES	(1) riskshare17h	(2) riskshare17h	(3) XTIVREG riskshare17h
Cognitive memory index[1]	0.025***	0.021**	0.238***
	(0.009)	(0.009)	(0.085)
Riskshare 2016	0.111***		0.114**
	(0.037)		(0.052)
Sex (Female dummy)		-0.060***	-0.013
		(0.021)	(0.033)
Age, years		-0.000	0.002
		(0.001)	(0.002)
Education, years		0.006**	0.000
		(0.003)	(0.004)
Farm size of parents		0.005	0.007
		(0.005)	(0.007)
Livestock endowment		0.003	0.019*
		(0.008)	(0.011)
Durable assets, number		-0.006	-0.032*
		(0.009)	(0.018)
Log(Income), 1000 ETB		0.007	0.027
		(0.010)	(0.017)
Idiosyncratic shock 2016-17, dummy			-0.053
			(0.033)
logconsinv1617			-0.001
			(0.004)
loganiminv1617			-0.002
			(0.004)
logprodassinv1617			-0.001
			(0.005)
logbusinv1617			-0.002
			(0.003)
Group fixed effects	Yes	Yes	Yes
Enumerator 2016 fixed effects	Yes	Yes	Yes
Enumerator 2017 fixed effects	Yes	Yes	Yes
Constant	0.367***	0.384***	-0.269
	(0.034)	(0.066)	(0.256)
Observations	938	930	929
R-squared	0.216	0.226	
Number of business groups	115	115	115

[1]Predicted in model (3), using lottery outcome in the 2016 game as instrument.
Log(Consumer goods investment+1)=logconsinv1617, Log(Livestock investment+1)=loganiminv1617, Log(Productive assets investment+1)=logprodassinv1617, Log(Business investment+1)=logbusinv1617, Log(Total investment+1)=logsuminv1617. Cluster-robust standard errors in parentheses, clustering on business groups, *** p<0.01, ** p<0.05, * p<0.1.

assumption that risk preferences are fairly stable and conclude that the changes in responses between the two game rounds are due to measurement error rather than preference change.

In summary, we learn three things from this. First, there exists substantial measurement error in the risky investment game as demonstrated by the weak correlation coefficient in the test-retest assessment. Second, the game is not a powerful tool to predict real world investment (main result). Third, risk tolerance and cognitive memory are positively related and this correlation may be hidden by endogeneity bias.

We will finally address the detected measurement error problem in the risk tolerance measure in a comprehensive way, building on Gillen et al. (2019). First we assess it for one investment variable at the time (equation (8)) using ORIV 2SLS models in Table 6. As a second approach to this, addressing that the different investment types are not independent, we estimate a stacked model for the four investment types, treating each of them as investment measures that are measured with error, see Table 7. These should give consistent estimates of the effects of risk tolerance, as proxied with the risky investment game, on investment.

As can be seen from Tables 6 and 7, the results from both approaches are disappointing. The coefficients on the predicted risk tolerance variable are large, vary in sign (Table 6), and have large standard errors. The statistical tests for endogeneity are unable to detect any significant endogeneity and the strength of the instrument across models is very weak. Unlike Gillen et al. (2019), who had a stronger correlation between the risk tolerance variables, our study reveals either that the measurement error is too high in our study for the risk tolerance measures to serve as good instruments for each other, or the risk tolerance variable is not endogenous, or both. We are tempted to believe that the limited asset integration in the game results in us not detecting any endogeneity. This implies that the game is unsuitable for obtaining a measure of broad bracketing risk tolerance for individuals for our fairly large sample. This may also indicate that there is limited within-subject correlation between myopic loss aversion, if this is what the game reveals (Gneezy and Potters, 1997), and real world investment behavior. However, more work is needed to assess the predictive power of alternative tools for eliciting risk tolerance, including dis-aggregated parameters based on RDU or PT.

Alternative simple experimental tools may be refined and assessed based on their predictive power. If substantial measurement error is a general characteristic of responses in hypothetical as well as incentivized experiments, there may be no way around collecting repeated measures for each subject and then to separate noise from these to get more reliable measures of individuals' risk preferences. Such tools must still ask questions that are simple to understand and should not require numeracy skills beyond what exist among the study subjects. Even the simplest tools may be fuzzy at first and the impulse response from the Type 1 ("Thinking Fast") system may be different from the responses that emerge after several questions with variation in the framing that may trigger more thinking and a switch to the Type 2 ("Thinking Slow") system of thinking that is more careful and calculated (Kahneman, 2017). This may also imply that decisions go from narrow bracketing to wider bracketing and thereby becoming more integrated with subjects' background wealth.

6. Conclusions

Our study assesses whether the incentivized risky investment game can provide a good measure of risk tolerance that can be used to predict subjects' real world investment propensity. For a fairly large sample of 966 (957) young business group members in 116 business groups in Ethiopia we first used the risky investment game integrated in a baseline survey. We then followed up with a new survey round one year later to collect investment data from the same subjects in combination with a hypothetical version of the game, after first recording their memory of the game played one year earlier. Our study reveals a low within-subject correlation of 0.135 for the investment level between the two game rounds although a better cognitive memory is associated with higher correlation but still being only 0.20 for those with perfect memory of their own decision and the outcome of the game one year earlier. Measurement error is therefore substantial in the game as the large changes in responses across the two game rounds cannot meaningfully be interpreted to be an indication of preference change. The true risk tolerance of the subjects is only partially affecting their decisions in each game round.

The second interesting finding in our study is that cognitive memory is positively correlated with the investment level in the game and more so than it looks like at first sight. After correcting for endogeneity bias in cognitive memory, subjects with better cognitive memory have a much stronger (10x) tendency to invest in the game.

Our results reveal that the investment behavior in the risky investment game is unable to predict the real world investment behavior of our subjects. We arrived at this conclusion after having used the ORIV approach of Gillen et al. (2019) to correct for measurement error and get consistent estimates of the risk tolerance effect on investment. This outcome could be due to high measurement error leading to weak correlation and thereby weak instruments. It could also be due to a fundamental problem with the risky investment game. We suggest this may be because the game, with its up-front allocation of cash, results in narrow bracketing as revealed by the lack of correlation with subjects'

Table 6
ORIV investment models for each investment type

VARIABLES	(1) log consinv1617	(2) log animinv1617	(3) log prodassinv1617	(4) log businv1617	(5) log suminv1617
Riskshare, predicted	-8.204	2.036	1.854	-4.855	-0.945
	(6.547)	(4.858)	(3.392)	(6.127)	(2.793)
Sex (Female dummy)	-1.159**	-1.089**	-0.207	-1.036**	-0.894***
	(0.539)	(0.434)	(0.298)	(0.499)	(0.265)
Age, years	-0.040*	-0.024	0.013	0.009	-0.000
	(0.021)	(0.020)	(0.012)	(0.022)	(0.010)
Education, years	0.170***	-0.034	-0.029	0.085*	0.016
	(0.047)	(0.042)	(0.030)	(0.048)	(0.025)
Farm size of parents	0.114*	0.041	0.057	0.029	0.006
	(0.068)	(0.060)	(0.044)	(0.073)	(0.037)
Livestock endowment	0.124	-0.039	0.056	0.058	0.001
	(0.119)	(0.113)	(0.078)	(0.138)	(0.061)
Durable assets, number	0.181	0.312**	0.261***	0.193	0.184***
	(0.138)	(0.127)	(0.084)	(0.161)	(0.070)
Log(Income), 1000 ETB	0.327**	0.493***	0.143	0.110	0.238***
	(0.148)	(0.143)	(0.098)	(0.171)	(0.082)
Idiosyncratic shock 2016-17, dummy	-0.169	-0.332	0.764***	-0.076	0.037
	(0.352)	(0.316)	(0.203)	(0.362)	(0.172)
Group fixed effects	Yes	Yes	Yes	Yes	Yes
Enumerator fixed effects	Yes	Yes	Yes	Yes	Yes
Constant	9.267***	5.787***	1.070	2.286	8.082***
	(3.106)	(2.155)	(1.960)	(3.104)	(1.288)
Observations	1,888	1,888	1,888	1,888	1,888
R-squared		0.263	0.376	0.116	0.311

[1] Instruments used for prediction: Riskshare17hyp for riskshare and vice versa.
Log(Consumer goods investment+1)=logconsinv1617, Log(Livestock investment+1)=loganiminv1617,
Log(Productive assets investment+1)=logprodassinv1617, Log(Business investment+1)=logbusinv1617,
Log(Total investment+1)=logsuminv1617.
Cluster-robust standard errors in parentheses, clustering on subjects, *** p<0.01, ** p<0.05, * p<0.1.

resource endowments. While the game has been used to study myopic loss aversion in the laboratory, our results may also indicate that real world investments are unaffected by myopic loss aversion in our study. While the game has been suggested used for the collection of data on risk tolerance in field contexts for its simplicity, we recommend testing whether other simple tools to reveal risk tolerance can do a better job at predicting real world investment behavior. More research is clearly needed as we are still at an early stage of developing good tools for field experiments within the area of risk and uncertainty.

A. Appendix 1: Experimental protocol and survey questions

The appendix contains survey questions and the experimental protocol (separate file).

CRediT authorship contribution statement

Stein T. Holden: Conceptualization of this study, Methodology, Training of field staff, Data checking and organization, Analysis, Write-up. **Mesfin Tilahun:** Training of field staff, Field testing, Fieldwork organization and supervision, Data checking and cleaning, Commenting on drafts.

Table 7
Stacked ORIV investment models combining the four investment types

VARIABLES	(1) invest	(2) invest
Riskshare, predicted	-5.713	-2.065
	(5.512)	(2.986)
Sex (Female dummy)		-0.853***
		(0.242)
Age, years		-0.012
		(0.010)
Education, years		0.047**
		(0.023)
Farm size of parents		0.062*
		(0.035)
Livestock endowment		0.043
		(0.063)
Durable assets, number		0.244***
		(0.074)
Log(Income), 1000 ETB		0.270***
		(0.079)
Idiosyncratic shock 2016-17, dummy		0.041
		(0.173)
Group fixed effects	Yes	Yes
Enumerator fixed effects	Yes	Yes
Constant	6.644***	4.567***
	(2.173)	(1.496)
Observations	7,676	7,604
R-squared	0.177	0.193

Cluster-robust standard errors in parentheses, clustering on subjects
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

References

Beauchamp, J.P., Cesarini, D., Johannesson, M., 2017. The psychometric and empirical properties of measures of risk preferences. Journal of Risk and Uncertainty 54, 203–237.

Binswanger, H.P., 1981. Attitudes toward risk: Theoretical implications of an experiment in rural india. Economic Journal 91, 867–890.

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

Charness, G., Viceisza, A., 2016. Three risk-elicitation methods in the field-evidence from rural senegal. Review of Behavioral Economics 3, 145–171.

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

Dasgupta, U., Mani, S., Sharma, S., Singhal, S., 2019. Can gender differences in distributional preferences explain gender gaps in competition? Journal of Economic Psychology 70, 1–11.

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

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

Gillen, B., Snowberg, E., Yariv, L., 2019. Experimenting with measurement error: Techniques with applications to the caltech cohort study. Journal of Political Economy 127, 1826–1863.

Gneezy, U., Leonard, K.L., List, J.A., 2009. Gender differences in competition: Evidence from a matrilineal and a patriarchal society. Econometrica 77, 1637–1664.

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

Holden, S.T., Quiggin, J., 2017. Climate risk and state-contingent technology adoption: shocks, drought tolerance and preferences. European Review of Agricultural Economics 44, 285–308.

Holden, S.T., Tilahun, M., 2018. The importance of ostrom's design principles: Youth group performance in northern ethiopia. World Development 104, 10–30.

Holden, S.T., Tilahun, M., 2020. Farm size and gender distribution of land: Evidence from ethiopian land registry data. World Development 130, 104926.

Holden, S.T., Tilahun, M., 2021. Endowment effects in the risky investment game? Theory and Decision , 1–16.

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

Kahneman, D., 2017. Thinking, fast and slow.

Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. Econometrica 47, 263–292.

Kimball, M.S., Sahm, C.R., Shapiro, M.D., 2008. Imputing risk tolerance from survey responses. Journal of the American statistical Association 103, 1028–1038.

Lilleholt, L., 2019. Cognitive ability and risk aversion: A systematic review and meta analysis. Judgment and Decision Making 14, 234–279.

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

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