Norwegian University of Life Sciences (NMBU)

How Large is the Endowment Effect in the Risky Investment Game?

Stein T. Holden and Mesfin Tilahun

Norwegian University of Life Sciences
Centre for Land Tenure Studies

Centre for Land Tenure Studies Working Paper 04/21
ISBN: 978-82-7490-295-4

ICAE-conference 2021 manuscript No.
(will be inserted by the editor)

How Large is the Endowment Effect in the Risky Investment Game?

Stein T. Holden · Mesfin Tilahun

Revised and presented: 30.08.2021

Abstract The risky investment game of Gneezy and Potters (1997) has been a popular tool used to estimate risk tolerance and myopic loss aversion. Holden and Tilahun (2021) tested and found that the simple one-shot version of this game that is attractive as a simple tool to elicit risk tolerance among respondents with limited education, produce significant endowment effects in two variants of the game where alternatively safe and risky initial monetary endowments are allocated. In this paper we use an alternative treatment that does not induce endowment effects. This allows us to establish a benchmark to assess the relative size of the endowment effects when initial safe and risky endowments are provided (contribution 1). While Prospect Theory could predict endowment effects in the game, it fails to explain the dominance of interior choices (partial investment). We propose an alternative endowment effect theory that gives predictions that are more consistent with the observed partial investment behavior (contribution 2).

Keywords Risky investment game · Endowment effects · Loss aversion · Utility curvature · Field experiment · Ethiopia

Stein T. Holden
School of Economics and Business, Norwegian University of Life Sciences, P. O Tel.: +47-94970515
E-mail: stein.holden@nmbu.no *Present address:* of F. Author

Mesfin Tilahun
School of Economics and Business, Norwegian University of Life Sciences, P. O,Box 5003, 1433 Ås, Norway and

Department of Economics, Mekelle University, Mekelle, Tigray, Ethiopia
E-mail: mesfin.tilahun.gelaye@nmbu.no

1 Introduction

The risky investment game was first introduced by Gneezy and Potters (1997) and was used to study myopic loss aversion based on Prospect Theory (PT). A one-shot version of the game was later introduced by Gneezy et al. (2009) and this simplified version has gained popularity as an easy tool to measure risk tolerance in the field (Charness and Viceisza, 2016). Here is a risk, however, that those using the game do not take into account that the game induces "endowment effects" associated with the initial allocation of money in the game. Endowment effects (Thaler, 1980) have usually been associated with commodities and not money. However, recent studies have revealed that safe and even risky monetary endowments may induce endowment effects (Bateman et al., 2005; Schmidt et al., 2008; Holden and Tilahun, 2021).

Holden and Tilahun (2021) showed that there was a substantial difference in the levels of investment when subjects were initially allocated a safe versus a risky initial monetary endowment and were freely allowed to trade between the risky and safe amounts and at the same constant exchange rate. The alternative safe and risky initial endowments created endowment effects in opposite directions.Their study, however, did not establish the benchmark that allowed them to identify the investment level when there is no endowment effect. We build on their study and our first contribution is to establish this benchmark which allows us to compare the relative size of the endowment effects for safe versus risky initial monetary endowments.

The dominant explanation for endowment effects has been based on PT and loss aversion that can be invoked in settings without risk as subjects are asked to give up endowments they possess (Thaler, 1980). The loss in utility associated with giving up one good is greater than the gain in utility from getting the same good.

The curvature of the utility or value function is a critical element to explain behavior in these games. PT usually assumes no asset integration and relates utility to deviations from a reference point. PT also usually assumes diminishing sensitivity around the reference point such that the value function is convex in the loss domain and concave in the gain domain. Full asset integration would imply risk neutrality for small gambles and no loss aversion and empirical studies provide strong evidence that full asset integration has to be rejected.The degree of kink in the value function at the reference point is a measure of the loss aversion that also may explain the endowment effect. A linear value function with a kink at the reference point will lead to either no or full investment in the risky investment game. The standard PT assumption with convex value function in the loss domain enhances the likelihood of corner solutions in the game. Based on PT it is therefore puzzling to find that interior choices dominate in several recent large sample studies (Dasgupta et al., 2019; Holden and Tilahun, 2021). E.g. Dasgupta et al. (2019) finds interior choices for more than 90% of a sample of 2000 Indian students. Our study, building on Holden and Tilahun (2021), similarly finds a strong dominance of interior choices (partial investment) among resource-poor young business group members in Ethiopia. We suggest that some theoretical modifications are needed to explain this. We provided a theoretical framework with some modifications of the standard PT and show that this framework provides predictions closer to the observed empirical results.

We compare the baseline T1 ("Safe Base") and the T2 ("Full risk") treatments of Holden and Tilahun (2021) with a new T3 ("Binary") treatment that does not invoke any endowment effects as no monetary endowments are provided during the game (only one is selected for real payout after all responses are given, as an incentive). In T1 the respondents are provided an initial endowment X=30 ETB[1], of which they are free to invest any amount $0 \leq x \leq X$ in a 50-50 lottery that will pay out $3x$ or 0 (Holden and Tilahun, 2021). This treatment is compared to the alternative treatment T2 ("Full Risk") where the respondents are initially provided the full 50-50 lottery of $3X$=90 ETB or 0 that they can sell themselves out of at the same exchange rate between $y = 3x$ between risky and sure money as in T1 (Holden and Tilahun, 2021). T3 includes a set of binary choices between combinations of risky and safe amounts with the same trade-off between these as in T1 and T2. If T1 and T2 create endowment effects in opposite directions (Holden and Tilahun, 2021), and T3 does not, T3 can provide a better basis for eliciting the utility curvature, given a functional form assumption such as CRRA. Rational subjects without a reference point bias should have the same investment levels across T1-T3.

The paper is organized as follows. Part 2 of the paper outlines the experimental design. Part 3 presents the theoretical framework and hypotheses. Part 4 describes the sample and implementation characteristics of the field experiment. Part 5 explains the estimation strategy. Part 6 presents the results and part 7 discusses the findings in relation to the relevant literature and part 8 concludes and makes some suggestions for further work.

2 Experimental design

The risky investment game is a real game with money. The one-shot version of the game was first used by Gneezy et al.(2009). The subjects are given an initial endowment, X=30 ETB. They can invest a share of this endowment, x/X (multiples of 5 ETB), in a 50-50 lottery with the outcome ($3x$) or 0. The unlucky ones remain with $X-x$ and the lucky winners with $X-x+3x = X+2x$ (treatment T1 (Safe base).

Holden and Tilahun (2021) introduced the risky initial monetary endowment, a 50-50 lottery prospect of $3X$=90 ETB or 0 (treatment T2 (Full risk)), which is the maximum risky investment level in treatment T1. The respondents were then offered to sell all or part of the lottery prospect and would then get a payment of one-third of the lottery winning value they would sell which is the same exchange rate as in treatment T1. If they sell y out of $3X$,

[1] ETB=Ethiopian Birr, 30 ETB was approximately a daily wage in the study area at the time of the study.

they will get $y/3$ as payment (multiples of (15,0) ETB for 5 ETB). Losers of the game will get $y/3$ and winners will get $3X - y + y/3$=$3X - 2y/3$.

The new treatment T3 does not offer any initial endowment and the reference point is the pre-game status quo. This treatment is implemented as a set of binary choices similar to the strategy method or the identification of a switch point in a Multiple Price List through repeated binary choices. No endowment is allocated to the respondents till after all binary choices have been made. The first binary choice is between getting X with certainty (no risky investment) and the $50 - 50$ lottery of getting $3X$ or 0 (full risk). The preferred choice in this first binary choice is then offered in the second binary choice versus the alternative choice that is a combination of $X/2$ for sure and a 50-50 lottery of $3X/2$ or 0, risking half of the safe amount (Expected value: $0.5 * 3X/2$, assuming $p(0.5) = 0.5$). Further binary choices combining safe and risky amounts, using the same exchange rate between these as in T1 and T2. This is done till an optimal mix of safe and lottery amounts is identified.

The sequence of binary choices in T3 allows "cross-over" in a way that can identify inconsistent responses. Such inconsistencies were only found for a small share of the respondents (4.1%). This indicates that inconsistent or random choices are not common among our sample of respondents. Treatments T1 and T2 are not capable of identifying such inconsistencies. This is thus an added strength of treatment T3. As T3 does not invoke endowment effects associated with an initial safe or risky monetary endowment, it does not invoke loss aversion and is therefore better suited to elicit information about the curvature of the utility function. A weakness of T3 is that it is more time consuming to implement.

Details of the experimental protocols (English version) for the three treatments are provided in Appendix 2. These were translated to the local language, Tigrinya, which was the language used in the field. The enumerators were trained with both versions and we ensured that the translations were accurate and that the enumerators understood the questions correctly and standardized the exact wording for all the questions and explanations.

3 Theory and hypotheses

It may not be obvious to the applied development economist that the one-shot risky investment game invokes loss aversion. S/he may therefore be tempted to frame the game within Expected Utility Theory (EUT) without integrating the game money with the wealth of the respondents. It is especially not common to assume that monetary endowments induce endowment effects due to loss aversion. EUT has for long dominated economic thinking related to risky choice among applied economists. Within the EUT framework risk preferences are captured by the utility curvature over the risky and safe amounts in the one-shot risky investment game:

$$\max EU(x) = 0.5u(30 - x) + 0.5u(30 + 2x) \tag{1}$$

A concave utility function is necessary to get interior solutions in the game for x with $0 < x^* < 30$. The optimal (preferred) level of $x_i = x_i^*$ for each subject i, is identified for each of the treatments. When imposing a specific functional form on the utility function such as a Constant Relative Risk Aversion (CRRA) function, assuming that there are no endowment effects, the relative risk aversion parameter (r) and its distribution in a sample population may be derived from the observed investment distribution.

With behavior according to EUT, a subject's allocation decisions should not vary across treatments in our experiment as behavior according to EUT implies no endowment effects (reference point bias) due to loss aversion or probability weighting:

$$EUT : x_i^*(T1) = x_i^*(T3) = x_i^*(T2) \tag{2}$$

Given a specific functional form of the utility function such as CRRA, a subject-specific individual risk aversion parameter can be derived based on her/his optimal x^* allocation that would be identical across the three treatments:

$$r_i^*(T1) = r_i^*(T3) = r_i^*(T2) \tag{3}$$

Given a CRRA-utility function, no asset integration, no endowment effect, and objective probability judgment, the relationship between r and optimal investment level is illustrated in Fig.1.

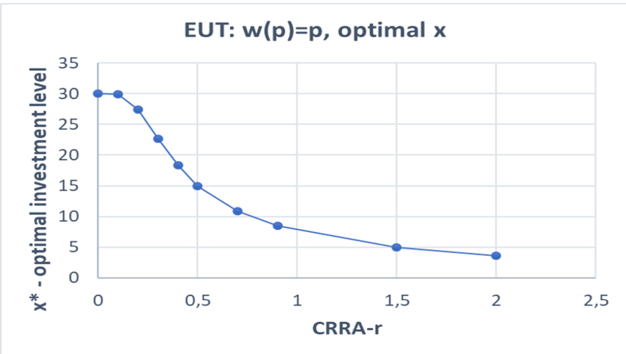

Fig. 1 EUT: CRRA-r and optimal investment in one-shot game

However, if real behavior deviates from EUT because of reference point effects, loss aversion and/or probability weighting, equations (2) and (3) will not hold and this would lead to biased estimates of risk aversion if wrongly imposing EUT when solving for r for T1 and T2. However, T3 should allow us to measure individual r_i.

Alternatively, the decisions in the game may be modeled based on the Prospect Theory (PT) to assess whether this theoretical framework is better as a basis to explain behavior across T1-T3 (Schmidt et al., 2008; Holden and

Tilahun, 2021). This model assumes the reference point is the initial monetary endowment in treatments T1 and T2 and it is only deviations from the reference point that matter. PT assumes diminishing sensitivity around the reference point, implying a convex value function in the loss domain and a concave value function in the gains domain around the reference point. Loss aversion is captured as a kink in the value function at the reference point. Assuming PT for T1 (Safe base), the reference point is the sure amount of 30. The decision-maker then maximizes the following expression (denoting loss aversion as λ):

$$\max PT(T1) = w^{+}(0.5)v(2x) - w^{-}(0.5)\lambda v(|x|) \quad (4)$$

For T2 (Full risk) it is the subjective value of the risky lottery yielding 90 with 0.5 probability which is the reference point, building on Holden and Tilahun (2021). This less salient (endogenous) reference point in T2 is denoted R. For T2 under PT the decision-maker seeks to maximize:

$$\max PT(T2) = w^{+}(0.5)v(90 - 2/3y - R) - w^{-}(0.5)\lambda v(|(y/3 - R|) \quad (5)$$

This model holds as long as $90 - 2/3y - R \geq 0$. Respondents will choose optimal y^* such that they avoid violation of this inequality. Given two respondents i, j with reference points $R_i > R_j$ who are identical in all other respects than their reference points, will choose optimal levels of y^* such that $y_i^* < y_j^*$. If T2 gives a higher reference point than T1 ($R > 30$), combined with loss aversion, the optimal investment level will be higher in T2 than in T1.

The T1 and T2 treatments alone do not allow us to identify the size of the endowment effect [2]. T1 and T2 pull in opposite directions compared to T3 which has no endowment effect. Holden and Tilahun (2021) demonstrated a substantial difference between the choices in T1 and T2. Building on the same data, we measure the size of the investment difference for T1 and T2 versus the new treatment T3.

Some recent findings in the experimental literature have found evidence of quite flat (close to linear) value functions under PT (Cheung, 2019). However, this is not consistent with what has been found in the one-shot risky investment game (Charness and Viceisza, 2016; Dasgupta et al., 2019; Holden and Tilahun, 2021). This may point towards a need to modify the theory or a need for an alternative theory or a need to modify or relax parts of PT, especially the assumption about the functional form of the value function in the loss domain. This may be done by replacing diminishing sensitivity with increasing or linear sensitivity in the loss domain. The functional form of the value function in the loss domain has not been much studied. Particularly among poor decision-makers living close to their survival constraint, such as our study subjects do, sensitivity could be increasing in the loss domain. There are also some studies that have found resource-poor subjects to have become less risk tolerant after

[2] The endowment effect could here be measured as the investment difference for T1 and T2 versus T3. We use investment shares, x/X and the deviations in these around $x(T3)/X$ in our analysis

a negative shock rather than becoming risk loving (Liebenehm et al. 2021; Holden and Quiggin 2017 +++ ref search).

It is also possible that $w^+(0.5) \neq 0.5$. Furthermore, we cannot be sure that the degree of loss aversion is as strong for the lottery prospect (T2) as for the sure amount in T1. It is also less obvious what the reference point is in T2, e.g. whether decision-makers apply a mean-variance perspective or use the maximum gain or loss as the reference point. It is also possible that subjects separate utility of safe and risky amounts before they aggregate them. Our experience is that this type of respondents with limited numeracy skills are not used to calculate average returns but rather use the more salient safe or risky amounts as reference points.

Based on this elaboration we propose two alternative endowment effect theory models (AEET1 and AEET2) where the respondents directly impose utility costs to the safe or risky prospects that they were endowed with in T1 and T2 and that they decide to trade in the game. The first of the two variants of the alternative theory uses standard probability weighted aggregation (AEET1), and the second aggregates separate utilities for safe and risky amounts (possibly allowing a preference for certainty) (AEET2). In these alternative models, T1 invokes an endowment effect when giving up safe amounts for risky amounts (a δ^s utility weight is associated with the safe endowment reduction). The AEET models allow for probability weighting like in prospect theory and rank-dependent utility theory. Unlike in prospect theory, the AEET models retain the concave utility in the loss domain. We do not rule out that giving up safe amounts (T1) can invoke a stronger endowment effect than giving up risky (lottery) amounts (T2), i.e. $\delta_s \geq \delta_r$. For AEET1, the (sophisticated with more numeracy skills) subjects maximize the following expression for T1:

$$\max AEET1(T1) = -\delta_s[(u^s(30) - u^s(30-x)] + [1 - w^+(0.5)]u^r(30-x) + w^+(0.5)u^r(30+2x) \quad (6)$$

Alternatively, with safe amounts being a focal point, possibly allowing for preference for certainty[3], subjects may distinguish between utility of certain amounts, $u^s(.)$, and utility of risky amounts, $u^r(.)$. The maximization problem can be reformulated as follows:

$$\max AEET2(T1) = -\delta_s[(u^s(30) - u^s(30-x)] + u^s(30-x) + w^+(0.5)[u^r(30+2x) - u^s(30-x)] \quad (7)$$

For T2, the endowment effect is associated with the giving up of (part of) the risky lottery opportunity when converting it to a safe amount. The sophisticated subjects maximize the following problem for T2:

$$\max AEET1(T2) = -\delta_r w^+(0.5)[u^r(90) - u^r(90-y)] + w^+(0.5)u^r(90-2y/3) + [1 - w^+(0.5)]u^s(y/3) \quad (8)$$

[3] Possibly captured by the probability weighting function or the utility function, or both.

With $y > 0$, the endowment effect is associated with the sacrificed opportunity to win $u^r(90)$ instead of $u^r(90 - y)$.

Alternatively, if utility of certain amounts is handled separately from utility of risky amounts, the model may be reformulated as follows:

$$\max AEET2(T2) = -\delta_r w^+(0.5)[u^r(90) - u^r(90 - y)] + w^+(0.5)u^r(90 - y) + u^s(y/3) \quad (9)$$

We do not believe in a single "correct" model of how people chose their reference points, or how they separate and aggregate utilities of risky and safe amounts. While a person with strong numeracy skills may argue that they make many mistakes when making these decisions, their decisions may still be their preferred choices based on the way they think and the expectations they have. Based on AEET1 and AEET2, we hypothesize that the optimal investment level, x^*, is different in the three treatments T1, T2 and T3 as follows:

$$x^*(T1) \leq x^*(T3) \leq x^*(T2) \quad (10)$$

We also hypothesize that the probability that a random respondent invests the full amount $x^* = X$ ($x = 30$ in T1 and $y = 0$ in T2) differs for the three treatments:

$$P(x^*(T1) = 30) \leq P(x^*(T3) = 30) \leq P(y^*(T2) = 0)|\delta_s, \delta_r > 0 \quad (11)$$

Overall, we assess whether EUT, PT or the AEET theories are better at describing the observed outcomes across treatments and subjects. EUT predicts no treatment effects and dominance of interior choices (solutions) and has already been rejected based on the finding of strong treatment effects. PT predicts significant treatment effects and dominance of corner solutions. We use AEET1 as the basis for running simulations based on a CRRA utility function and moderate levels of dis-utility associated with giving up safe and risky endowments. We demonstrate that this model predicts significant treatment effects and high likelihood of interior solutions while the level of investment varies with r where r takes on a reasonable range values.

4 Sampling and implementation

The starting point is the same data as used by Holden and Tilahun (2021) for treatments T1 and T2 and that come from a field experiment with rural youth business group members in northern Ethiopia. Treatment T3 was first pilot tested together with T2 and then scaled up to a much larger sample among the same population in 2019. These were land-poor rural youth and young adults that due to their poverty had been found eligible to join youth business groups as an alternative source of livelihood in their home communities (*tabias*). Their average age was 31 years and with a standard deviation of 10 years. The mean level of education was five years, varying from no education to 12 years of completed education. Financial and business skills are important for them to succeed in their business activities. Men dominated in the groups and constituted close to 70 percent of the group members.

Treatment T1 (Safe initial endowment) was used in a baseline survey in the study area in 2016 for a sample of 1138 youth business group members in 119 business groups in five districts in the Tigray region of Ethiopia. Treatments T2 (Full risk) and T3 (Binary) were implemented in 2019, first in a pilot experiment (N=243 for T2 and N=304 for T3), and then treatment T3 was scaled up to a larger sample of youth business group members from the same districts (N=2184) as for treatment T1 in 2016. Treatments T2 and T3 were randomized at group level for the sample of youth business groups and group members in the pilot district. A large share of the sample in the 2019 pilot experiment also participated in the 2016 experiment, thereby facilitating a combination of a within-subject and between-subject design.

The initial endowment of 30 ETB used as the safe amount was equivalent to a daily rural wage rate in agriculture in the study areas in 2016. For practical reasons the investment levels were allowed to be 0, 5, 10, 15, 20, 25 and 30 ETB. Further splitting into a finer sub-division would require the use of coins which we wanted to avoid.

Local schools were used as field labs. One youth group was interviewed at a time with 12 enumerators doing the experiments and interviews of 12 members simultaneously. Three classrooms were used, locating an experimental enumerator and a group member in each corner of a classroom. This prevented communication between group members during the games. It also implied that the enumerators never interviewed or did experiments with more than one group member per group, thereby ensuring orthogonality between groups and enumerators, to control for and minimize potential enumerator bias. Payouts for the experiments took place immediately after completion of the interviews within a group.

5 Estimation strategy

The share invested from the maximum safe amount ($X = 30$ ETB) is used as the measure of the risky investment level. This implies that $r = \frac{x}{X}$ and $0 \leq r \leq 1$.

We use the risky investment share as a dependent variable and start with parsimonious linear panel data models that include all treatments from the 2016 and 2019 rounds for the full sample, including the pilot district. District fixed effects and enumerator fixed effects were included as controls.

To assess the relative size of the endowment effects in treatments T1 and T2, we have included treatment T3 which is not invoking any endowment effects. We estimated linear panel data models with variants of the following specification to compare the sizes of the endowment effects in terms of changing the investment share levels in T1 and T2 compared to T3:

$$r_{gi} = r_1 + \alpha_2 T1_g + \alpha_3 T2_g + \alpha_{4d} D_d + \alpha_{5e} E_d + \alpha_{gs} s_{gi} + g_g + \epsilon_{gi} \quad (12)$$

Subscript g represents group, subscript i represents individual, r_1 represents the estimated share invested in treatment T3 (with no endowment effect). α_2 captures the endowment effect for T1 with T3 (no endowment effect) as baseline investment share level, and is expected to be negative. α_3 represents the endowment effect for treatment T2 and is expected to be positive if risky endowments induce endowment effects. D_d represents a vector of district dummy variables, E_d represents a vector of enumerator dummy variables, s_{gi} represents a set of individual characteristics (sex, age, birth rank, education), g_g represents group random effects, and ϵ_{gi} represents the error term.

The pilot experiment in one district in 2019 allowed a direct identification of the endowment effect in the full risk (T2) treatment. It could also measure the endowment effect for treatment T1 by combining these data with the data from 2016 for the same district. We imposed a number of robustness checks to assess the stability of the treatment effect, including community, group and individual fixed effects. We also investigated whether changes in age and shock exposures in the period 2016-19 period could explain changes in the responses in the game and imposed controls for such possible effects, see Appendix A2.

The initial tests for the robustness of the results in the full sample included the addition of individual controls (gender, age, birth rank and education). Another potential source of bias could be the enumerators used in the experiments. While they were doing only one interview per group each, we had a change in enumerators from 2016 to 2019 based on the quality of their work and availability (selection of the best available ones for the 2019 survey and dropping some poor performers). The inclusion of enumerator fixed effects controls for such possible enumerator bias. We had five enumerators that participated in both years and as an additional robustness check we run a separate model for the sample of enumerators that were involved in both years to assess whether that change in enumerators from 2016 to 2019 could lead to selection bias (model (3) in Table 3). We refer to Appendix A2 for additional robustness checks.

6 Results

Fig. 2 shows the full sample investment distribution for all three treatments. The figure illustrates highly significant differences in distributions across the three treatments. Fig. A1 in the Appendix shows the risky investment distribution for treatment T1, comparing the pilot district (Degua Tembien) distribution with that of the full sample. Degua Tembien was the district where the pilot test of treatments T2 and T3 took place in 2019. It can be seen that the response distribution in the pilot district is very similar to that in the full sample. Figure 2 shows the distribution of investments in the pilot district in 2019 for T2 (Full Risk) and T3 (Binary) (243 versus 304 respondents). We see that a substantially larger share invested the full amount in T2 than in T3. We attribute this difference to the endowment effect in treatment T2. However, interior choices dominate in all three treatments. This indicates that the utility function must be non-linear, implying that loss aversion alone combined with a linear utility function cannot explain the investment levels for most respondents in the game.

Table 1 presents average shares invested out of the maximum safe amount that can be obtained for the three treatments in the full sample and in the pilot district. Table 2 shows the shares investing the full amount by treatment in the full sample. Table 3 assesses the statistical significance of the treatments using Wilcoxon ranksum/Mann-Whitney tests for the shares invested by sample type. The test results demonstrate highly significant treatment effects ($p < 0.01$) for all treatment differences, except in the same enumerator sample where the sample size gets very small for the T2 (Full Risk) sample and $p = 0.15$).

Table 4 presents the results from linear panel data models with youth group random effects, district fixed effects, and enumerator fixed effects and with standard errors corrected for clustering at the youth group level. Treatment T1 (Safe base) serves as the baseline treatment in all regression models and its investment share is captured by the constants in the tables. However, we need to remember that T1 has an endowment effect and that the coefficient for treatment T3 is due to the endowment effect associated with treatment T1. Models (1), (2) and (3) are for the full sample. Model (2) includes additional individual controls and Model (3) includes treatment and gender interactions. Model (4) includes the sample for which the same enumerators were used in 2016 and 2019 as an extra robustness check for potential enumerator selection bias.

Table 5 presents models for the pilot district, combining the 2016 and 2019 data and imposing alternative controls for unobserved heterogeneity. Model (5) includes group random effects, Model (6) includes group fixed effects and Model (7) includes individual fixed effects. All the models include enumerator fixed effects.

The main findings from the experiments are as follows:

Result 1: *Treatment T2 results in a significantly higher average investment level than treatment T3 and a larger share of respondents that invests the full amount than treatment T3.*

Result 1 indicates that there is a significant endowment effect associated with the allocation of the risky amount in treatment T2.

Result 2: *Treatment T1 (Safe base) resulted in a significantly lower average investment level than T3 (Binary) and gives a share investing the full amount that is significantly lower than T3.*

Result 2 indicates that T1 induces a significant endowment effect that pulls in opposite direction compared to T2.

Tables 4 demonstrates that the treatment effects are robust to the inclusion of additional controls. The individual control variables were also assessed for their systematic variation across treatments, see Appendix Table A1. As treatment T1 was implemented in 2016 it is not surprising to find a significant

Table 1 Mean shares invested by treatment and sample

— Full sample —				
Treatment	Mean	Median	St.Err	N
T1:Safe Base	0.443	0.333	0.007	1138
T2:Full Risk	0.691	0.833	0.021	243
T3:Binary	0.565	0.667	0.007	2184
— Pilot district —				
Treatment	Mean	Median	St.Err	N
T1:Safe Base	0.425	0.333	0.015	249
T2:Full Risk	0.691	0.833	0.021	243
T3:Binary	0.611	0.667	0.019	330
— Same enumerators —				
Treatment	Mean	Median	St.Err	N
T1:Safe Base	0.460	0.333	0.011	487
T2:Full Risk	0.609	0.667	0.035	102
T3:Binary	0.560	0.500	0.011	898

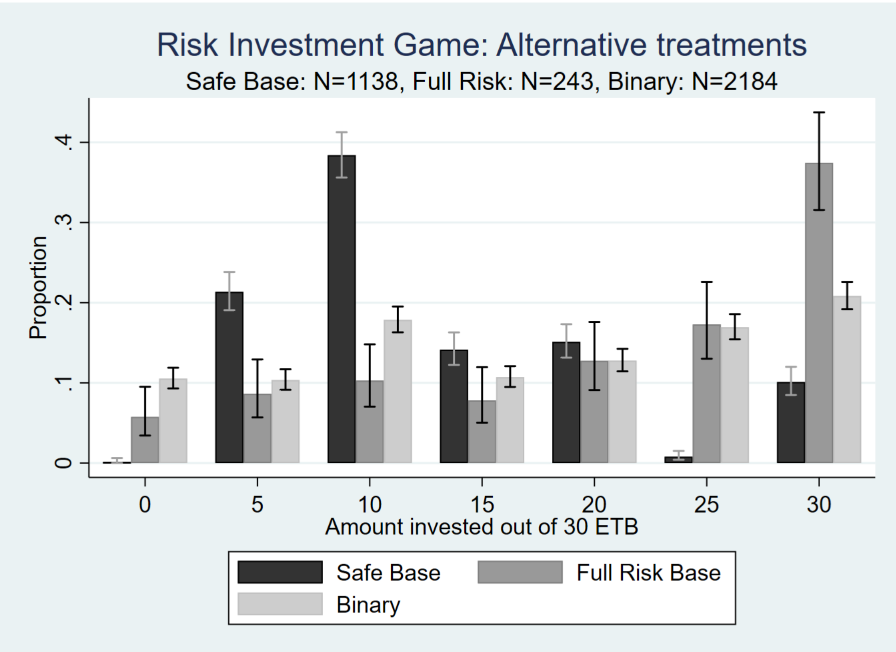

Fig. 2 Distribution of investments in Treatments 1, 2 and 3 (full sample)

Table 2 Share investing full amount by treatment

Treatment	Mean	St.err.	N
T1 (Safe Base)	0.101	0.009	1138
T2 (Full Risk)	0.374	0.031	243
T3 (Binary)	0.208	0.009	2184
All	0.185	0.007	3565

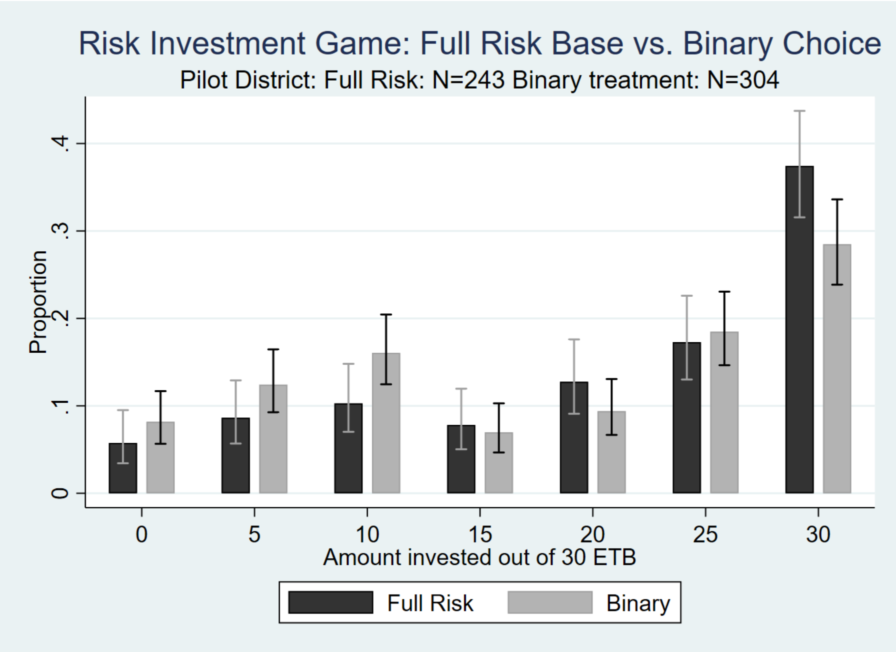

Fig. 3 Robustness check Treatment 2 (Full Risk) and Treatment 3 (Binary) in pilot district

Table 3 Wilcoxon-Mann-Whitney tests by treatment and sample

	Full sample		Degua Tembien		Same enumerator	
	z-score	P-value	z-score	P-value	z-score	P-value
T1 vs. T2	-10.965	0.0000	-9.078	0.0000	-3.993	0.0001
T2 vs. T3	5.744	0.0000	2.770	0.0056	1.425	0.1542
T1 vs. T3	-10.487	0.0000	-6.448	0.0000	-5.321	0.0000

age difference between T1 versus T2 and T3. Age had, however, very limited effect on the investment levels as can be seen in Table 5. Age is insignificant in Model (2) and significant at 5 percent level in Model (3) but with a very low coefficient. Five years higher age is associated with a 1 percentage point lower investment share. The difference in age cannot therefore explain the large differences in investment levels between T1 versus T2 and T3. The age effect even points in opposite direction of the change in mean investment levels in 2016 compared to 2019, when the group members have become three years older.

To further inspect the robustness of the results, the pilot district sample is used with alternative controls, see models (3) - (5) in Table 5. We utilize the fact that for this district many of the same youth groups and group members were included in the 2016 as well as 2019 samples. This allows us to impose stronger controls for unobserved time-invariant heterogeneity through the use of group fixed effects and individual fixed effects. We see from Table 5 that the

Table 4 Full sample and same enumerator models with controls

VARIABLES	(1) Full sample	(2) Full sample	(3) Same enumerators
T1 treatment	-0.096***	-0.101***	-0.111***
	(0.017)	(0.018)	(0.019)
T2 treatment	0.110***	0.110***	0.050
	(0.027)	(0.025)	(0.039)
Male		0.046***	0.054***
		(0.012)	(0.018)
Age of member		-0.000	-0.002**
		(0.001)	(0.001)
Birth rank		0.005**	0.005
		(0.002)	(0.004)
Education, years		0.006***	0.002
		(0.001)	(0.002)
Constant	0.472***	0.415***	0.507***
	(0.021)	(0.033)	(0.047)
Observations	3,564	3,564	1,487
Number of youth groups	308	308	305

All models with district FE and enumerator FE
T3 is baseline treatment (Constant)
Cluster-robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

treatment effects were robust to these alternative specifications. Treatment T1 gives investment levels that are robust and about 10 percentage points lower than that of treatment T3. Treatment T2 gives somewhat smaller treatment effects in the range 6-9.8 percentage points higher than that of T3. While the endowment effect for T1 is slightly larger than for T2, this difference is not statistically significant. Both risky and safe endowments are therefore associated with substantial endowment effects.

To assess whether shocks could contribute to the changes between 2016 and 2019 (T1 versus T2 and T3) we ran robustness checks for the pilot district as well as the full sample where we included a dummy variable for whether respondents had been exposed to any shocks during the last 12 months before the 2019 experiments and survey. The empirical evidence is mixed on whether risk preferences are influenced by shocks, with some studies finding that risk preferences are stable (Sahm, 2012), some studies finding that shocks make subjects less risk tolerant (Cassar et al., 2017; Liebenehm, 2018), and yet other studies finding that shocks make subjects more risk tolerant (Hanaoka et al., 2015; Voors et al., 2012). Some studies find that idiosyncratic shocks do not affect risk preferences while covariate shocks do, but deviations from this finding are also found (Liebenehm, 2018). Our variable captured idiosyncratic shocks like serious sickness or death in the family, violence, crime exposures, and production losses due to unfavorable weather. The results from these tests are included in Appendix 1, Tables A3 and A4. The shock variable was insignificant in all models. This indicates that the changes from treatment T1

Table 5 Robustness checks for pilot district (Degua Tembien)

VARIABLES	(1)	(2)	(3)	(4)	(5)
T1 treatment	-0.105***	-0.102***	-0.105***	-0.100***	-0.114*
	(0.033)	(0.034)	(0.032)	(0.033)	(0.064)
T2 treatment	0.081***	0.095**	0.083***	0.098**	0.060
	(0.029)	(0.047)	(0.030)	(0.047)	(0.049)
Male, dummy			0.041*	0.035	
			(0.025)	(0.027)	
Age of member			-0.001	-0.000	
			(0.001)	(0.002)	
Birth rank			0.009	0.008	
			(0.006)	(0.007)	
Education, years			0.006*	0.006	
			(0.004)	(0.004)	
Group RE	Yes	No	Yes	No	No
Group FE	No	Yes	No	Yes	No
Individual FE	No	No	No	No	Yes
Constant	0.533***	0.532***	0.463***	0.456***	0.550***
	(0.045)	(0.043)	(0.074)	(0.084)	(0.064)
Observations	822	822	822	822	822
R-squared		0.141		0.149	0.292
Number of youth groups	53	53	53	53	
Number of yg members					593

All models with enumerator FE.
Cluster-robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

in 2016 to T2 and T3 in 2019 cannot be explained by such shocks affecting the respondents and their responses from 2016 to 2019. We refer to the Appendix for all the robustness checks.

7 Discussion

We have introduced an alternative binary treatment approach to the one-shot risky investment game and proposed that this approach does not induce any endowment effect, unlike the standard one-shot version of the game (Gneezy et al., 2009) and the risky base treatment introduced by Holden and Tilahun (2021). We therefore find endowment effects for money, including lottery money and that the endowment effect for lottery money is almost as large as that for safe money. Our study is in a rural economy where cash is scarce and this could potentially enhance the endowment effect for money.

The one-shot risky investment game can easily be incorporated in large sample surveys and more easily so than the more complicated Multiple Price List approaches that may be more cognitively demanding to respond to. Holden and Tilahun (2021) showed that the game is associated with significant endowment effects and that EUT should not be used to estimate utility curvature in form of a single parameter based on the game results. However, they did not show how large the relative bias of such a parameter is in the case of safe and risky initial parameters. The new binary treatment introduced here does not initiate an endowment effect and is therefore more suited for the elicitation of the utility curvature as well as to get measures of the relative size of the endowment effects in the safe and risky initial amount versions of the one-shot risky investment game.

While Prospect Theory can explain and predict different behavior across treatments, it does not explain the dominance of interior choices in the game under the assumption of diminishing sensitivity in the value function. Loss aversion in combination with a linear value function would also lead to "bang-bang" (corner) solutions. This is far from what we observe for all three treatments as the share of interior solutions was close to 0.9 for T1, about 0.58 for T2 and close to 0.68 for T3. Theoretical modifications may therefore be needed to explain the dominance of interior solutions. We suggested two alternative endowment effect models (AEET1 and AEET2) and where AEET1 requires somewhat stronger numeracy skills than AEET2. We use simple simulations below to assess the ability of AEET1 to predict interior solutions and variation across the three treatments where we attribute the variation in the investment levels to the variation in utility curvature.

We use a CRRA utility function and with an endowment effect $\delta = 0.1$ for treatments T1 and T2 and compare with T3 (new "control"). Fig. 4 shows optimal x^* for alternative values of utility curvature parameter r for the three treatments with $\delta = 0.1$.

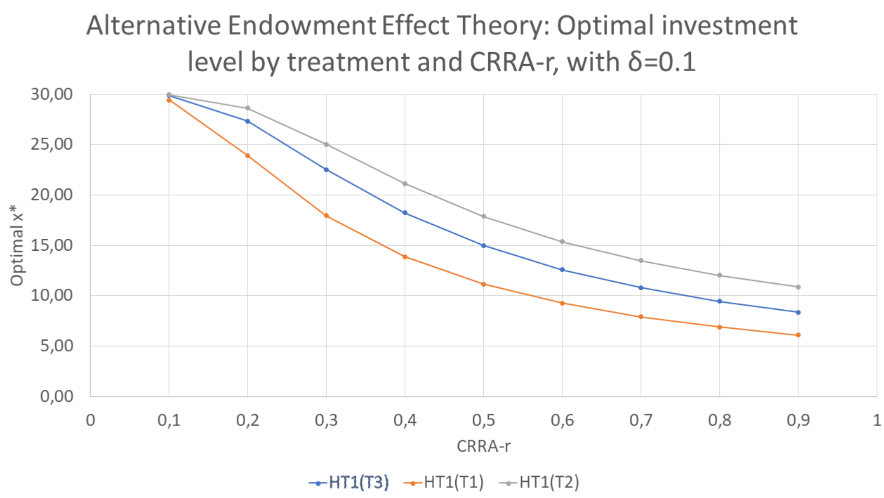

Fig. 4 Optimal investment (x*) by treatment and CRRA-r, with endowment effect

Fig. 4 illustrates that for the AEET1 model interior solutions dominate for a wide range of r values. The choice of an endowment effect parameter $\delta = 0.1$ creates treatment differences close to the average treatment differences observed in the data. This implies an endowment (utility) effect of about 10% of the utility of money or lottery value given up in T1 and T2. This illustrates an alternative way of modelling endowment effects than the kinked value function in prospect theory. This theory indicates that initial endowments received and"given up" in the experiment are not treated fully as sunk costs by the subjects.

8 Conclusion

The one-shot version of the risky investment game has gained popularity and has been proposed as particularly useful in field settings for respondents with limited numeracy skills (Charness and Viceizca 2016). Holden and Tilahun (2021) demonstrated that the game is associated with substantial endowment effects but did not assess the relative size of the endowment effects associated with safe and risky amounts. In this paper we introduce an alternative treatment that allow us to investigate the relative size of the endowment effects for safe and risky amounts of money allocated in the game. We find that both safe and risky amounts are associated with substantial endowment effects and that the endowment effect for risky money is (almost) as large as that for safe money that are initially provided in the game.

We also found that interior choices dominated in all three treatments in the game while Prospect Theory, based on the diminishing sensitivity around the reference point assumption, predicts"all or nothing" decisions in the game. We have proposed an Alternative Endowment Effects Theory (AEET) and demonstrate with simple simulations that it predicts the dominance of interior solutions and that a reasonable endowment effect parameter can predict the observed treatment effects. We conclude that the binary version of the game (T3) can be used to estimate utility curvature as long as the probability weighting function $w(0.5) = 0.5$ is approximately correct on average. One drawback of the binary choice approach is that it is somewhat more cumbersome to introduce than the simple one-shot game version. We recommend further testing of variants of it that also may be used to inspect the consistency of within-subject responses, tests that are not feasible with the simple one-shot version. We consider the development of experimental tools for elicitation of risk preferences in populations with limited education and numeracy skills to still be in its infancy. University students are not the best"test subjects" for such tools. Yet it is such populations that face the most serious risks as they are the most vulnerable and live in very risky environments. It is of high policy relevance to design simple tools that are capable of predicting their behavioral responses to shocks and risky or uncertain events. The risky investment game is a relevant candidate in this context and deserves further testing and scrutiny.

9 Acknowledgement

We acknowledge financial support from the Norwegian Agency for Development Cooperation (NORAD), the NORHED program for capacity-building and the Research Council of Norway for funding under the NORGLOBAL2 research program. We also acknowledge valuable comments to an earlier version of the paper from Dag Einar Sommervoll, Ferdinan Vieider, participants in the 42nd Annual Meeting of the Norwegian Association of Economists, NMBU, Ås, 6-7th January 2020, and participants in the ICAE 2021 digital conference on 30th August 2021.

References

Bateman, I., Kahneman, D., Munro, A., Starmer, C., and Sugden, R. (2005). Testing competing models of loss aversion: An adversarial collaboration. *Journal of Public Economics*, 89(8):1561–1580.

Cassar, A., Healy, A., and Von Kessler, C. (2017). Trust, risk, and time preferences after a natural disaster: experimental evidence from thailand. *World Development*, 94:90–105.

Charness, G. and Viceisza, A. (2016). Three risk-elicitation methods in the field-evidence from rural senegal. *Review of Behavioral Economics*, 3(2):145–171.

Cheung, S. L. (2019). Eliciting utility curvature in time preference. *Experimental Economics*, pages 1–33.

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

Gneezy, U., Leonard, K. L., and List, J. A. (2009). Gender differences in competition: Evidence from a matrilineal and a patriarchal society. *Econometrica*, 77(5):1637–1664.

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

Hanaoka, C., Shigeoka, H., and Watanabe, Y. (2015). Do risk preferences change? evidence from panel data before and after the great east japan earthquake. Technical report, National Bureau of Economic Research.

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

Liebenehm, S. (2018). Temporal stability of risk attitudes and the impact of adverse shocks—a panel data analysis from thailand and vietnam. *World Development*, 102:262–274.

Sahm, C. R. (2012). How much does risk tolerance change? *The Quarterly Journal of Finance*, 2(04):1250020.

Schmidt, U., Starmer, C., and Sugden, R. (2008). Third-generation prospect theory. *Journal of Risk and Uncertainty*, 36(3):203.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, 1(1):39–60.

Voors, M. J., Nillesen, E. E., Verwimp, P., Bulte, E. H., Lensink, R., and Van Soest, D. P. (2012). Violent conflict and behavior: a field experiment in burundi. *American Economic Review*, 102(2):941–64.

A Appendix 1.

A.1 Pilot district representativeness and individual controls

Fig. A1 assesses how representative the pilot district is compared to the full sample in terms of the distribution of responses in the risk investment game in treatment T1 in 2016.

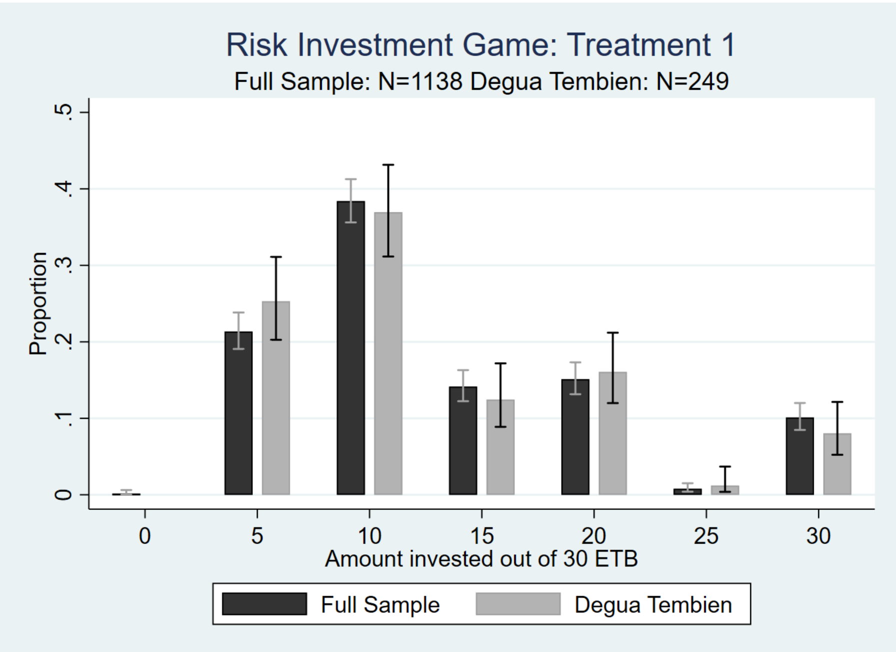

Fig A1. Pilot district vs. full sample for treatment T1

Table A1 presents the individual control variables and tests for their difference across treatments. With treatments T2 and T3 implemented three years later, it is expected that the respondents on average will be about three years older in these treatments.

A.2 Robustness checks and individual controls

We ran joint regression models combining the 2016 and 2019 data for the pilot district. Additional specifications with group fixed effects and individual fixed effects were included as the groups and individuals for this district to a large extent overlapped in the two years.

Table A1. Individual characteristics by treatment: t-tests

	T1 Safe Base	T2 Full Risk	T3 Binary	t-tests T1 vs T2	t-tests T1 vs. T3	t-tests T2 vs. T3
Age, years	29.07	32.78	32.24	-3.710***	-3.170***	0.540
	(9.796)	(9.216)	(9.507)	(0.685)	(0.351)	(0.641)
Birth rank	3.105	3.198	3.37	-0.093	-0.265***	-0.172
	(2.002)	(1.877)	(2.183)	(0.140)	(0.078)	(0.146)
Education, years	5.345	5.078	4.608	0.267	0.737***	0.470
	(3.978)	(3.747)	(3.968)	(0.278)	(0.145)	(0.267)
Observations	1138	243	2184	1381	3322	2427

This allowed control for time-invariant group and individual unobservable characteristics (see Appendix Table A4).

Our design confounds year with the T1 treatment effect and there is a risk that the youth have changed their behavior in the baseline treatment over this three years period. We scrutinize this in the following ways:

a) By including individual characteristics (sex, age, birth rank and education) and inspect whether the gain in age over the three years could have changed their responses (Appendix Table A1 assesses differences in the individual characteristics across treatments (and years for treatments T2 and T3 versus treatment T1.

b) By including an individual level shock dummy variable for those that had experienced a serious shock over the last 12 months before the 2019 round. The shock variable included individual and family health shocks, death in the family, climate, violence, crime and other shocks. The shock variable was included as an additional control in the full sample as well as in the pilot district sample models (model results in Appendix Tables A3 (full sample) and A4 (pilot district)).

The linear panel data models yield coefficients that are marginal effects and are convenient to interpret for that reason. Since our dependent variable is a share with values from zero to one, we also estimated fractional probit models that take this into account. We have not included the results from these models, however, because they gave marginal effects that were very close to those from the linear panel data models.

Table A3. Have recent shocks influenced the responses?

	(1)	(2)	(3)
VARIABLES			
Sample	Full	Full	Same enumerators
T1 treatment	-0.096***	-0.101***	-0.112***
	(0.017)	(0.017)	(0.019)
T2 treatment	0.110***	0.110***	0.051
	(0.027)	(0.027)	(0.044)
2018-19 Shock dummy	-0.008	-0.003	-0.016
	(0.021)	(0.022)	(0.042)
Male, dummy		0.046***	0.054***
		(0.012)	(0.019)
Age of member		-0.000	-0.002**
		(0.001)	(0.001)
Birth rank		0.006**	0.005
		(0.002)	(0.004)
Education, years		0.006***	0.002
		(0.002)	(0.003)
Constant	0.473***	0.416***	0.507***
	(0.021)	(0.032)	(0.049)
Observations	3,564	3,564	1,487
Number of youth groups	308	308	305

All models with enumerator FE
Cluster robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A4. Have recent shocks in the pilot district affected the results?

	(1)	(2)	(3)
VARIABLES			
Panel controls	Group RE	Group FE	Individual FE
T1 treatment	-0.102***	-0.101**	-0.117*
	(0.039)	(0.047)	(0.067)
T2 treatment	0.080**	0.092*	0.064
	(0.032)	(0.055)	(0.055)
2018-19 Shock dummy	0.032	0.031	-0.043
	(0.043)	(0.045)	(0.097)
Constant	0.528***	0.528***	0.529***
	(0.042)	(0.045)	(0.080)
Observations	822	822	822
R-squared		0.141	0.292
Number of youth groups	53	53	
Number of yg members			593

All models with enumerator FE
Cluster robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix 1.
Risky Investment Game: Treatment T3: Instructions

This is an experiment with alternative lotteries and sure money where you have to make choices depending on your preferences. The lotteries involve a 50-50 chance of winning. You will respond to a number of paired alternatives that are more or less risky and where you choose the one you prefer of each of the two prospects. After you have responded to all binary choice options, one of the alternatives will be chosen and played for real (lottery). You may win up to 90 ETB. Whether you win or not is determined by throwing a 20-sided die. If it lands on numbers 1-10, you lose and get nothing of the lottery amount, only the safe amount. If it lands on numbers 11-20 you win the full amount in the lottery.

18	Do you agree to play the game? 1=Yes, 2=No	Code	
19a	First choice: What do you prefer of these two alternatives? 1. 50% chance of winning 90 ETB, or 2. Get 30 ETB for sure If choice 1, go to **19b**. If choice 2, go **19c**	Code	
19b	What do you prefer? 1. 50% chance of winning 90 ETB, or 2. 15 ETB for sure AND 50% chance of winning 45 ETB? If choice 1, go to **19d**. If choice 2, go **19e**	Code	
19c	What do you prefer? 1. Get 30 ETB for sure, or 2. 25 ETB for sure AND 50% chance of winning 15 ETB? If choice 1, go to **End of experiment (payout)**. If choice 2, go to **19h**	Code	
19d	What do you prefer? 1. 50% chance of winning 90 ETB, or 2. 5 ETB for sure AND 50% chance of winning 75 ETB? If choice 1, go to **20 (Lottery + payout)**. If choice 2, go **19e**	Code	
19e	What do you prefer? 1. 5 ETB for sure AND 50% chance of winning 75 ETB? 2. 10 ETB for sure AND 50% chance of winning 60 ETB? If choice 1, go to **20 (Lottery + payout)**. If choice 2, go 19f	Code	
19f	From 19c. What do you prefer? 1. 10 ETB for sure AND 50% chance of winning 60 ETB? 2. 15 ETB for sure AND 50% chance of winning 45 ETB? If choice 1, go to **20 (Lottery + payout)**. If choice 2, go to **19g**	Code	
19g	What do you prefer? 1. 15 ETB for sure AND 50% chance of winning 45 ETB? 2. 20 ETB for sure AND 50% chance of winning 30 ETB? If choice 1, go to **20 (Lottery + payout)**. If choice 2, go to **19h**	Code	
19h	What do you prefer? 1. 25 ETB for sure AND 50% chance of winning 15 ETB? 2. 20 ETB for sure AND 50% chance of winning 30 ETB? If choice 1, go to **20 (Lottery + payout)**. If choice 2, go to **20 (Lottery + payout)**.	Code	
Information after all paired choices have been made: The last preferred choice will be subject to the lottery (unless you preferred the safe amount and no lottery). The interviewer then plays the lottery with you for the remaining lottery amount **with the die where numbers 1-10 imply loss and numbers 11-20 imply that you win**.			
20	Outcome of lottery, 1=Win, 0=Loss	Code	