



Endowment effects in the risky investment game?

Stein T. Holden¹  · Mesfin Tilahun^{2,3}

Accepted: 30 April 2021
© The Author(s) 2021

Abstract

The risky investment game of Gneezy and Potters (Q J Econ 112(2):631–645, 1997) has been proposed as a simple tool to measure risk aversion in applied settings, especially attractive in settings where participants may have limited education. However, this game can produce a significant endowment effect (attached to the initial position), so that analysis of the behavior in this game should not be done in the Expected Utility Theory (EUT) framework. The paper illustrates this point, by showing that risk tolerance can be much higher when the initial endowment concerns a risky lottery.

Keywords Risky investment game · Endowment effects · Loss aversion · Salient reference points · Field experiment · Ethiopia

1 Introduction

The term “endowment effect” was first used by Thaler (1980) and he related this effect to the fact that losses are weighted more heavily than gains and associated this with prospect theory (PT) and loss aversion in settings without risk. The loss in utility associated with giving up one good is greater than the gain in utility from getting the same good; “losses loom larger than gains”. Third-Generation Prospect Theory (PT3) (Schmidt et al. 2008) provides the basis for endowment effects existing for monetary endowments and risky and uncertain prospects such as lottery tickets. Our study investigates whether there can be significant endowment effects in

✉ Stein T. Holden
stein.holden@nmbu.no

Mesfin Tilahun
mesfin.tilahun.gelaye@nmbu.no

¹ School of Economics and Business, Norwegian University of Life Sciences, Ås, Norway

² School of Economics and Business, Norwegian University of Life Sciences, 1433, P. O. Box 5003, Ås, Norway

³ Department of Economics, Mekelle University, Mekelle, Tigray, Ethiopia

the one-shot risky investment game (Gneezy et al. 2009), that built on the dynamic version of the game (Gneezy and Potters 1997). While the purpose of the dynamic version of the game was to illustrate myopia due to loss aversion and thereby endowment effects, it is less obvious that the one-shot version of the game has the same effect.

The beauty of the one-shot game is its simplicity and it has been proposed as particularly useful in field settings for respondents with limited numeracy skills (Charness and Viceisza 2016). An applied development economist may easily ignore the potential endowment effect associated with loss aversion in the game and frame the one-shot game results with Expected Utility Theory (EUT). While most studies using the risky investment game have not used the game to estimate a parameter for risk aversion based on EUT (capturing the curvature of the utility function) of respondents, this may be tempting and there are a few studies that also do this (e.g., (Crosetto and Filippin 2016; Dasgupta et al. 2019). If endowment effects/loss aversion plays a role in the one-shot game, they get biased estimates of the utility curvature.

Our experiment was implemented as a field experiment in rural Ethiopia with young business group members as subjects. These subjects had on average 6 years of education and were computer illiterate. The experiments were introduced to them by trained experimental enumerators using classrooms in local schools as “labs”.

Our “lab-in-the-field” experiment was designed to allow us to distinguish between safe and risky money that were provided as initial endowments. We assess whether an endowment effect is found for safe initial money provided as well as for an initial 50–50 lottery or whether the first only or none of these monetary endowments are associated with an endowment effect. In our baseline treatment T1 (“Safe Base”), the respondents were provided an initial endowment, of which they were free to invest any share of the initial endowment in a 50–50 lottery that would pay out three times the invested amount or nothing. This treatment is compared to an alternative treatment T2 (“Full Risk”) where the respondents were initially provided the full 50–50 lottery of three times the safe amount provided in treatment T1. In T2, they could sell themselves out of the risky lottery at the same exchange rate between risky and sure money as in T1.

Our results show a highly significant treatment effect demonstrating endowment effects and reference dependence in the game. EUT should therefore not be used to estimate risk aversion in the narrow sense of utility curvature based on the game.

The paper is organized as follows. Section 2 of the paper outlines the experimental treatments and theoretical framework. Section 3 describes the experimental procedures of the field experiment. Section 3 explains the estimation strategy. Section 4 presents the results and Sect. 5 discusses the findings in view of the alternative theories and Sect. 6 concludes and makes some suggestions for further work.

2 Experimental treatments: EUT vs. PT

2.1 Experimental treatments

The baseline treatment (T1) was based on the one-shot version of the risky investment game first used by Gneezy et al. (2009). Respondents were told that they would play a real game with money. In this game, they could choose to keep or invest the whole or part of an initial endowment $X = 30$ ETB.¹ They could invest a share x/X (multiples of 5 ETB) in a 50–50 lottery with the outcome $3x$ or 0. In case of loss, the respondent only received $X - x$. The lucky winners obtained $X - x + 3x = X + 2x$.

Treatment T2 endowed the respondents with a 50–50 lottery of $3X = 90$ ETB or 0, which was the maximum risky investment level in T1. The respondents were then offered to sell all or part of the lottery and would then get a payment of one-third of the lottery winning value they would sell. If they sold y out of $3X$, they would get $y/3$ as payment (y was in multiples of 15, so that the safe amounts received were in multiples of 5). Losers of the game would get $y/3$ and winners would get $3X - y + y/3 = 3X - 2/3y$.

Details of the experimental protocols (English version) for the 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 used the same wording in the local language for all the questions and explanations.

2.2 EUT vs. PT

To an applied development economist, it may not be obvious that the one-shot risky investment game invokes loss aversion. S/he may therefore interpret the experimental results through the lens of Expected Utility Theory (EUT). 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 under narrow bracketing,² risk preferences are captured by the utility curvature over the risky and safe amounts in the one-shot risky investment game [treatment T1 (safe base)]

$$\max EUT(x) = 0.5u(30 - x) + 0.5u(30 + 2x). \quad (1)$$

Risk aversion, captured by the concavity of the utility function, is necessary to get interior solutions for x with $0 \leq x^* \leq 30$. The optimal level of $x_i = x_i^*$ for each subject i is identified with the standard experiment. When imposing a specific functional form on the utility function such as a Constant Relative Risk Aversion (CRRA) function, the relative risk aversion parameter (r) and its distribution in a

¹ Ethiopian Birr (ETB) is the monetary unit in Ethiopia where the field experiment was run. 30 ETB was approximately a rural daily wage in 2016 in the study areas.

² This means that we assume no integration of experimental money with background wealth.

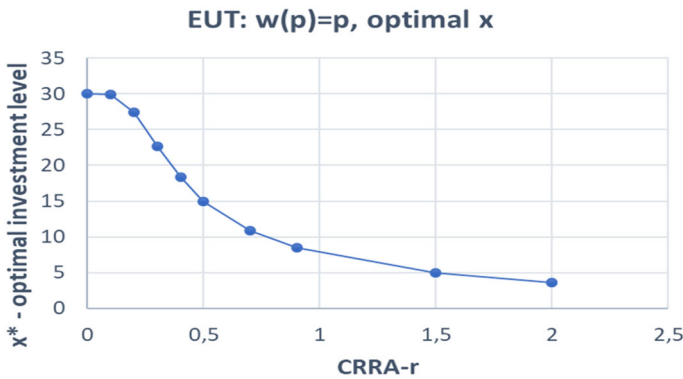


Fig. 1 EU: CRRA-r and optimal investment in one-shot game

sample population may be derived from the observed investment distribution based on the one-shot standard game, see Fig. 1.

For treatment T2, the EUT maximization problem can be stated as

$$\max EUT(y) = 0.5u(90 - (2/3)y) + 0.5u(y/3). \quad (2)$$

With behavior according to EUT, subject's allocation decisions should not vary across treatments in our experiment as behavior according to EU implies no endowment effects (reference point bias) due to loss aversion or probability weighting. This may be verified as the relationship between x in Eq. (1) and y in Eq. (2) is $x = 30 - y/3$ or $y = 90 - 3x$. By plugging in for x in Eq. (2), we see that it becomes identical to Eq. (1). The optimal investment level will be the same across the two treatments for a subject i behaving according to EUT:

$$EUT : x_i^*(T1) = x_i^*(T2). \quad (3)$$

Given a specific functional form of the utility function such as CRRA, this, therefore, leads to the same individual risk aversion (utility curvature) parameter derived for each subject based on her/his optimal x^* allocation that would be identical across the two treatments. Using the one-shot game to measure risk aversion would then lead to no bias in the estimation of risk aversion. Given a CRRA-utility function, no integration of prospect money with background wealth, no endowment effect, and objective probability judgment, the relationship between CRRA-r and optimal investment level is illustrated in Fig.A1 in Appendix 1.³

However, if real behavior deviates from EUT because of reference point effects, loss aversion, and/or probability weighting, the conversion between x and y and Eq. (3) will not hold and using T1 based on EUT to estimate the utility curvature as the measure of risk aversion would lead to biased estimates.

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

³ Some include background consumption such as a daily wage when estimating risk aversion within EUT, see, e.g., Andersen et al. (2008).

explain behavior across T1–T2. One may even need inspiration from Third-Generation Prospect Theory (PT3) to think of there being endowment effects associated with monetary prospects in the one-shot risky investment game (Schmidt et al. 2008).⁴ This model assumes that the reference point is the endowment at the decision point in treatments T1 and T2 and it is only deviations from the reference point endowment that matter. PT typically 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. Loss aversion is represented 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. We denote this (endogenous) reference point R . 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 game does not allow us to identify respondents' reference points in T2 or how these reference points are associated with $w(0.5)$ and λ . However, a significantly higher investment level in T2 than in T1 allows us to reject EUT and is a clear indication of significant endowment effects in the game. Interior solutions in the game are also an indication of non-linear value functions.

The standard one-shot game has been shown to give significant gender differences with women investing significantly less than men in most earlier studies (Charness and Gneezy 2012). It is still a mystery why this game tends to give stronger gender differences than other games used to investigate gender differences in risk preferences (Filippin and Crosetto 2016). We ask whether this could be associated with an endowment effect bias that may be stronger for women. We test whether the gender difference is stronger in T1 than in T2, or whether it is eliminated or goes even in the opposite direction in T2. If the gender difference remains strong and in the same direction in all three treatments, we interpret this as an indication that women are less risk tolerant than men but not more loss averse.

⁴ We are thankful to a reviewer for proposing this model.

3 Experimental procedure

The respondents in the experiment were sampled from rural youth business groups in northern Ethiopia. The group members were resource-poor rural youth and young adults that due to their poverty had been found eligible to join youth business groups in their home communities (*tabias*) based on their land poverty, residence, and demonstrated interest in developing a rural livelihood in their home community. The average age was 31 years and with a standard deviation of 10 years, giving more age variation than the typical student samples used in laboratory experiments. The mean level of education was 5 years, but it varied from no education to 12 years of completed education. Still, financial and business skills are important for them to succeed in their business activities. Women constituted close to one-third of the group members.

Treatment T1 (safe base) was used in a baseline survey in the study area in 2016 for a sample of 1138 business group members in 119 business groups in five districts in the Tigray region of Ethiopia.

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. This was also the reason for multiplying the invested amount with three rather than the 2.5 factor used in the initial Gneezy and Potters (1997) study and several other studies.

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 in the estimation. Payouts for the experiments took place immediately after the completion of the interviews.

The low share of respondents investing the full amount in the 2016 experiment led the authors to worry that the design could lead to bias and reveal respondents as less risk tolerant than they really were. With new funding from a new project, a follow-up survey was planned in 2019. To test the hypothesis of an endowment effect in the game, T2 (full risk) was implemented as a pilot test in one of the districts.

A large share of the sample in this pilot study also participated in the 2016 experiment, thereby facilitating a combination of a within-subject and between-subject design. Treatment T2 was randomized at group level for the sample of youth business groups and group members ($N = 243$ for T2) in the pilot district.

4 Estimation strategy

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

We use the risky investment share as a dependent variable and start with parsimonious linear panel data models that include treatments T1 and T2 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 test the robustness of the treatment effects and to control for other variables, we estimated linear panel data models with variants of the following specification:

$$s_{gi} = \alpha_1 + \alpha_2 \text{Fullrisk}_g + \alpha_{3d} D_d + \alpha_{54} E_d + \alpha_{gs} z_{gi} + g_g + \epsilon_{gi}, \quad (6)$$

where subscript g represents group, subscript i represents individual, α_1 represents the estimated share invested in the baseline treatment (T1), α_2 captures the treatment effect for T2 as the mark-up share invested in the risky lottery, D_d represents a vector of district dummy variables, E_d represents a vector of enumerator dummy variables, z_{gi} represents a set of individual characteristics (sex, age, birth rank, education), g_g represents group random (fixed) effects, and ϵ_{gi} represents the error term.

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). To test the gender effect across treatments, we included the $\text{sex} * \text{T2}$ interaction term in one specification [model (3) in Table 2]. We furthermore carried out separate regressions for the pilot district (Degua Tembien) where treatment T3 was implemented. In Appendix 1, we also compare the response distribution under treatment T1 in the pilot district versus the full sample.

A 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 (included in all models). 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 (4) in Table 2]. We refer to Appendix A2 for additional robustness checks.

Shock exposure in the period between 2016 and 2019 could also have affected the preferences of the respondents and they have become 3 years older. We, therefore, assessed whether age and an individual risk exposure recall variable were correlated with the treatment effects. Models with the risk exposure variable are included in Appendix 1, Tables A2 and A3. The significance of the age variable can be inspected in all models where the socio-economic variables were included.

5 Results

Table 1 presents average shares invested out of the maximum safe amount that can be obtained for the two treatments in the full sample, in the pilot district, and for the same enumerator sample. Table 1 also shows the shares investing the full amount by treatment in the full sample. The table includes test results for the statistical significance of the treatments using Wilcoxon rank-sum/Mann–Whitney tests for the shares invested by sample type. The test results demonstrate a highly significant treatment effect ($p < 0.01$).

Figure 2 in Appendix 1 shows the full sample investment distribution for the two treatments. The figure illustrates highly significant differences in distributions. Figure 3 in Appendix 1 shows the 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 treatment T2 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. We see from Fig. 2 that a substantially larger share invested the full amount in T2 than in T1. Interior solutions dominate to a larger extent in T1.

Table 2 presents the regression results from linear panel data models for the full sample and the same enumerator sample 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 1 (Safe base) serves as the baseline treatment in all regression models and its investment share is captured by the constants in the tables. Models (1), (2), and (3) are for the full sample. Model (2) includes additional individual controls and Model (3) includes a treatment and gender interaction variable. 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 3 presents models for the pilot district, combining the 2016 and 2019 data and imposing alternative controls for unobserved heterogeneity. Model (1) includes group random effects, Model (2) includes group fixed effects, Model (3) includes group random effects and individual controls, and Model (4) includes group fixed effects and individual controls. All the models include enumerator fixed effects.

The main findings from the experiments are as follows:

Result 1: *Treatment T2 results in significantly higher average investment level and a much larger share of respondents that invests the full amount than for treatment T1.*

Result 1 indicates that there are significant endowment effects and EUT has to be rejected. The high share of interior solutions also indicates that the value function is non-linear for most respondents.

Table 2 demonstrates that the treatment effect is 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 age difference between T1 versus T2. Age had, however, a very limited effect on the investment

Table 1 Investment (shares) by treatment and sample

Treatment	Full sample			
	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
Treatment	Pilot district (Degua Tembien)			
	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
Treatment	Same enumerators			
	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
Treatment	Share investing full amount			St.err.
	Mean			
T1: safe base	0.101			0.009
T2: full risk	0.374			0.031
Sample	Wilcoxon–Mann–Whitney tests			
	z-score		P value	
Full sample	– 10.965		0.0000	
Pilot district (Degua Tembien)	– 9.078		0.0000	
Same enumerator	– 3.993		0.0001	

levels as can be seen in Tables 2, 3, 5, and 6. Age is insignificant in all models and with a parameter value no larger than 0.001. The age difference cannot, therefore, explain the large differences in investment levels between T1 and T2.

Result 2: *The investment level for women was lower than that of men in both treatments.*

The average gender difference was smaller for T2 than for T1, but the gender effect was not significantly different across the two treatments ($\text{Chi}2(1)=0.78$, $p=0.376$). A gender difference in loss aversion should result in an effect in opposite direction for T2 than for T1 and cannot, therefore, explain the strong gender difference found in the risky investment game in our sample.

To further inspect the robustness of the results, the pilot district sample is used without and with individual controls, see Table 3. We utilize the fact that for this district many of the same youth groups were included in 2016 as well as 2019 samples. This allows us to impose stronger controls for unobserved time-invariant

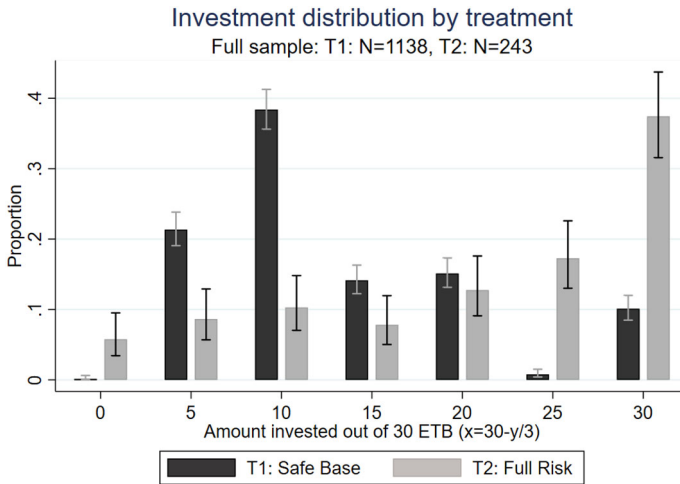


Fig. 2 Distribution of investments in Treatments 1 and 2)

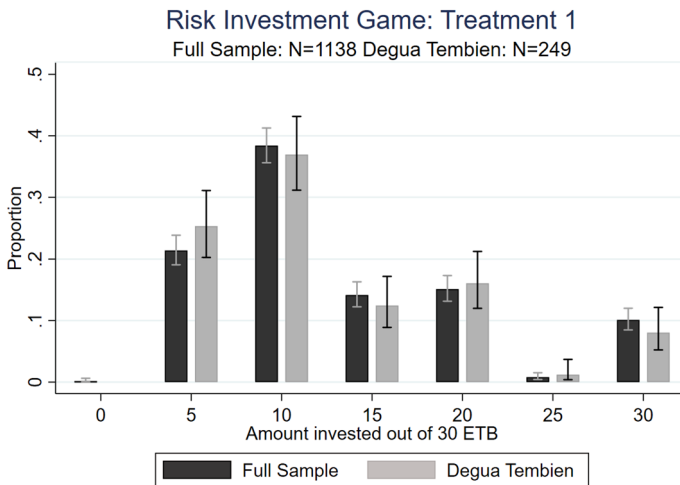


Fig. 3 Robustness check (Treatment 1): pilot district vs full sample

heterogeneity through the use of group fixed effects in addition to group random effects. We see from Table 3 that the full risk treatment (T2) effect is robust to these alternative specifications. T2 gives a significantly larger average investment level than T1 in all model specifications and the investment level is 15.7–16.9% points higher for T2 than for T1.

To assess whether shocks could contribute to the changes between 2016 and 2019 (T1 versus T2), 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

Table 2 Full sample and same enumerator models with controls

Variables	(1) Full sample	(2) Full sample	(3) Full sample	(4) Same enumerators
T2-full risk (female)	0.170*** (0.039)	0.161*** (0.040)	0.185*** (0.042)	0.148*** (0.043)
Male		0.058*** (0.015)	0.065*** (0.016)	0.053** (0.025)
T2-Full risk*Male			0.215*** (0.047)	
Age		0.001 (0.001)	0.001 (0.001)	0.000 (0.002)
Birth rank		0.002 (0.003)	0.002 (0.003)	0.005 (0.006)
Education (years)		0.007*** (0.002)	0.007*** (0.002)	0.005 (0.003)
Constant	0.389*** (0.026)	0.297*** (0.040)	0.293*** (0.041)	0.309*** (0.067)
Observations	1,381	1,381	1,381	589
Number of groups	129	129	129	129

With enumerator and district FE (not shown)

T1 (safe base) is baseline treatment (constant)

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

survey. The 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 A2 and A3. The shock variable was insignificant in all models. This indicates that the changes from treatment T1 in 2016 to T2 in 2019 cannot be explained by such recent shocks affecting the subjects and changing their responses from 2016 to 2019.

6 Discussion

Our experiment demonstrates not-so-obvious endowment effects in the standard one-shot risky investment game. It demonstrates even less obvious endowment effects for risky lottery allocations. We, therefore, find endowment effects for money, including lottery money. Our study is in a rural economy where cash is scarce and this could potentially enhance the endowment effect for money.

Endowing the respondents with the lottery treatment T2 dramatically increased the share of the respondents that accepted the full lottery (from 10.1 to 37.4% of the sample) as well as the average share invested in the game (from 44.3 to 69.1%) and the median investment level even more (from 33.3 to 83.3%) for the full sample.

Table 3 Robustness checks for pilot district (Degua Tembien) sample

Variables	(1) riskshare Group RE	(2) riskshare Group FE	(3) riskshare Group RE	(4) riskshare Group FE
T2-full risk	0.157*** (0.043)	0.169*** (0.051)	0.161*** (0.044)	0.163*** (0.053)
Male			0.056* (0.029)	0.031 (0.034)
Age			- 0.001 (0.001)	0.001 (0.002)
Birth rank			0.008 (0.007)	0.007 (0.007)
Education (years)			0.008* (0.004)	0.007 (0.005)
Constant	0.379*** (0.049)	0.379*** (0.045)	0.317*** (0.076)	0.273*** (0.080)
Observations	492	492	492	492
R-squared		0.203		0.213
Number of groups	34	34	34	34

All models with enumerator FE. Cluster-robust standard errors in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

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. Given its tractability for field experiments, it is of general interest to know whether the game can be used to generate simple and reliable estimates of relative risk aversion based on EUT. An implication of our study is that using EUT to predict a risk aversion parameter yields higher risk aversion (utility curvature) for the standard one-shot version of the game (T1) than for treatment T2. Using Fig. 1 to predict CRRA- r based on EUT for average investment shares gives $r = 0.57$ for T1 and $r = 0.34$ for T2. Close to 60% of the subjects invested 10 ETB or less in T1 and this could, according to EUT, imply that the majority of the sample has $r \geq 0.75$, while this applies to less than 25% of the sample for T2. One, therefore, has to treat investment levels in the game cautiously and take into account that the risk tolerance in the game is influenced by endowment effects due to loss aversion. The game does not allow a separation of the effects of utility curvature, loss aversion, and probability weighting that explain risk preferences under PT. While the standard one-shot game has the initial endowment as the obvious reference point, treatment T2 has a less salient reference point, but the experimental findings indicate that on average the respondents perceived the reference point in T2 to be higher than in T1 and this contributed to higher investment levels in T2.

Prospect theory (PT) can explain why investment levels are higher in T2 than in T1 as loss aversion plays a role. However, loss aversion cannot alone explain the dominance of interior solutions (90% and 57% interior solutions in T1 and T2, respectively) in the game. As the game allows no flexibility in objective probabilities, probability weighting can neither, together with loss aversion, explain the dominance of interior solutions. Non-linear value functions are therefore necessary to explain the dominance of interior solutions. This finding is somewhat puzzling as some recent studies have found close to linear value functions based on PT in experiments eliciting value function curvature, probability weighting functions, and loss aversion (Abdellaoui et al. 2013; Cheung 2019).

7 Conclusion

The one-shot version of the (Gneezy and Potters 1997) risky investment game has gained popularity and has been proposed as particularly useful in field settings for respondents with limited numeracy skills (Charness and Viceisza 2016). We have investigated whether the one-shot version of the game can invoke loss aversion and thereby endowment effects and thereby affect investment levels. Viewing the game with the lens of Expected Utility Theory can, if that is the case, lead to biased estimates of risk aversion (=utility curvature) as this theory ignores reference dependence. We found evidence of substantial endowment effects in the game demonstrating that risk behavior depends on (endogenous) reference points of the respondents. EUT is therefore not suitable for modeling behavior in the game. Finally, while the risky investment game has been found to give larger gender differences in risk tolerance than some other experimental approaches used to elicit risk preferences, we found that this gender difference was not explained by a gender difference in loss aversion and endowment effects.

A Appendix 1

A.1 EUT: CRRA-r and investment share

Figure 1 shows the relationship between invested amount and utility curvature (CRRA-r) under EUT under the assumption that the investment level is determined by the utility curvature.

A.2 Individual controls vs treatment

Table 4 presents the individual control variables and tests for their difference across treatments. With treatment T2 implemented 3 years later than treatment T1, it is expected that the respondents on average will be about 3 years older in treatments T2.

Table 4 Individual characteristics by treatment: *t* tests

	T1 Safe base	T2 Full risk	<i>t</i> tests T1 vs T2
Age, years	29.07 (9.796)	32.78 (9.216)	-3.710*** (0.685)
Birth rank	3.105 (2.002)	3.198 (1.877)	-0.093 (0.140)
Education, years	5.345 (3.978)	5.078 (3.747)	0.267 (0.278)
Observations	1138	243	1381

Table 5 Robustness check: recent shock effect in pilot district

Variables	(1) riskshare Group RE	(2) riskshare Group FE
Panel controls		
T2-full risk	0.161*** (0.047)	0.166*** (0.064)
2018-19 Shock dummy	- 0.008 (0.065)	- 0.024 (0.064)
Male	0.055* (0.030)	0.031 (0.034)
Age	- 0.001 (0.001)	0.001 (0.002)
Birth rank	0.008 (0.007)	0.006 (0.007)
Education (years)	0.008* (0.004)	0.007 (0.005)
Constant	0.315*** (0.076)	0.272*** (0.081)
Observations	492	492
R-squared		0.213
Number of groups	34	34
Number of subjects	376	376

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

A.3 Investment distribution by treatment

Figure 2 shows the distribution of invested amounts by treatment in the full sample.

Table 6 Robustness check: recent shock variable included

Variables	(1) riskshare Full	(2) riskshare Full	(3) riskshare Same enumerators
T2-Full Risk	0.172*** (0.043)	0.163*** (0.043)	0.160*** (0.048)
2018-19 Shock dummy	- 0.020 (0.063)	- 0.017 (0.063)	- 0.102 (0.100)
Male		0.058*** (0.016)	0.051** (0.025)
Age		0.001 (0.001)	0.000 (0.002)
Birth rank		0.002 (0.003)	0.005 (0.006)
Education (years)		0.007*** (0.002)	0.005 (0.003)
Constant	0.390*** (0.026)	0.298*** (0.040)	0.312*** (0.068)
Observations	1,381	1,381	589
Number of groups	129	129	129

Model (3) Same enumerators in 2016 and 2019

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

A.4 Representativeness of pilot district

Figure 3 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. There are no significant differences between the pilot district and the full sample.

A.5 Robustness checks for shock effects

Table 5 provides a robustness test for the pilot district sample and Table 6 provides a similar test for the full sample and the same enumerator sample. The shock variable is insignificant in all models.

Acknowledgements 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 earlier versions of the paper from Dag Einar Sommervoll, Ferdinand Vieider, participants in the 42nd Annual Meeting of the Norwegian Association of Economists, NMBU, Ås, 6–7th January 2020, and two anonymous reviewers.

Funding Open access funding provided by Norwegian University of Life Sciences.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as

you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abdellaoui, M., Bleichrodt, H., l'Haridon, O., & Paraschiv, C. (2013). Is there one unifying concept of utility? An experimental comparison of utility under risk and utility over time. *Management Science*, *59*(9), 2153–2169.
- Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2008). Eliciting risk and time preferences. *Econometrica*, *76*(3), 583–618.
- Charness, G., & Gneezy, U. (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior & Organization*, *83*, 50–58.
- Charness, G., & 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*, 1–33.
- Crosetto, P., & Filippin, A. (2016). A theoretical and experimental appraisal of four risk elicitation methods. *Experimental Economics*, *19*(3), 613–641.
- 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.
- Filippin, A., & Crosetto, P. (2016). A reconsideration of gender differences in risk attitudes. *Management Science*, *62*(11), 3138–3160.
- Gneezy, U., Leonard, K. L., & List, J. A. (2009). Gender differences in competition: Evidence from a matrilineal and a patriarchal society. *Econometrica*, *77*(5), 1637–1664.
- Gneezy, U., & Potters, J. (1997). An experiment on risk taking and evaluation periods. *The Quarterly Journal of Economics*, *112*(2), 631–645.
- Schmidt, U., Starmer, C., & 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.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.