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

by Stein T. Holden and Mesfin Tilahun

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

Submitted: 09.04.2021

Abstract The risky investment game of Gneezy and Potters (1997) has been a popular tool used to estimate risk tolerance and myopic loss aversion. We have assessed whether a simple one-shot version of this game that is attractive as a simple tool to elicit risk tolerance among respondents with limited education, can produce significant endowment effects associates with the investment from the initial monetary endowment allocated. Holden and Tilahun (2021) test whether a similar endowment effect can be found in relation to the allocation of a risky monetary prospect in a field experiment and find highly significant endowment effects. In this paper we use an alternative treatment that should not impose 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 first (contribution 1). While Prospect Theory could predict endowment effects in the game, it fails to explain the dominance of interior solutions in all treatments. We propose an alternative endowment effect theory that gives predictions that are more consistent with observed behavior (contribution 2).

Keywords Risky investment game · Endowment effects · Loss aversion · Salient reference points · 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 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 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”. There could be other explanations than loss aversion for this kind of ”exchange asymmetries”(Plott and Zeiler, 2007), or “stickiness” of endowments and these have been described with different names such as “status quo bias” (Samuelson and Zeckhauser, 1988; Kahneman et al., 1991), “anchoring effects” (Epley and Gilovich, 2001; Simonson and Drolet, 2004), and ”default effects” or ”default pull effects”(Cappelletti et al., 2014; Dhingra et al., 2012; Sunstein, 2002). Marzilli Ericson and Fuster (2011) disentangle endowment effects from ownership and show that expectations affect reference points and may thereby trigger endowment effects. Third Generation Prospect Theory (PT3) (Schmidt et al., 2008) which provides the basis for endowment effects existing for monetary endowments and for risky and uncertain prospects such as lottery tickets. Bateman et al. (2005) tested and found such endowment effects for money as well as commodities and explained them as being due to loss aversion.

Our experiment is based on the one-shot version of the risky investment game of Gneezy and Potters (1997), first used by Gneezy et al. (2009). While the purpose of the dynamic version of the game was to create myopic loss aversion and thereby endowment effects, it is not obvious that the one-shot version of the game has the same effect. This is what we aim to investigate. An applied development economist who ignores this, may combine the one-shot game results with Expected Utility Theory and attribute investment levels in the game to the curvature of the utility function as a measure of risk aversion. If endowment effects/loss aversion play a role in the one-shot game, s/he gets a biased estimate of the utility curvature.[1]

In the baseline treatment T1 (”Safe Base”) the respondents are provided an initial endowment X=30 ETB, 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). In this paper we compare these two treatments with treatment T3 (”Binary”), where no initial endowment is provided. 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.

For T3 the reference point is the pre-game status quo as we use the strategy method and allocate no endowments in the sequence of binary choices. If T1 and T2 create endowment effects in opposite directions (Holden and Tilahun,

[1] We define risk aversion narrowly as utility curvature in the sense of EUT.

2021), and T3 does not, T3 can provide a better basis for eliciting the utility curvature, given a functional form assumption such as CRRA.

Holden and Tilahun (2021) find highly significant and substantial allocation differences between T1 and T2 but cannot identify the relative size difference in these endowment effects as they pull in opposite directions. Safe endowments (T1) may create stronger endowment effects than risky endowments (T2). Furthermore, the reference point is less salient in the risky endowment treatment than the safe endowment treatment.

They find that corner solutions (no investment or full investment are dominated by interior solutions in the game. Prospect Theory with loss aversion and a linear utility function cannot explain the dominance of interior solutions, which points towards non-linear utility in this small stakes risky game. We propose an alternative theory to better predict the dominance of interior solutions and the size of the endowment effects in treatments T1 and T2.

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 baseline treatment (T1) was based on the one-shot version of the risky investment game first used by Gneezy et al.(2009). Respondents are told that they will play a real game with money. In this game they can choose to keep or invest the whole or part of an initial endowment X=30 ETB. They can 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 receives $X - x$. The lucky winners obtain $X - x + 3x = X + 2x$.

Holden and Tilahun (2021) introduced treatment T2 where the respondents are offered a 50-50 lottery prospect of $3X$=90 ETB or 0, which is the maximum risky investment level in 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. If they sell y out of $3X$, 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 $3X - y + y/3$.

The new treatment T3 does not offer any initial endowment and the reference point should be 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 and $50 - 50$ lottery of getting $3X$ or 0. 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 (Expected value: $0.5 * 3X/2$). Further binary choices are provided till an optimal mix of safe and lottery amounts is identified. 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 used the same exact wording in the local language for all the questions and explanations.

3 Theory and hypotheses

It is not obvious to the applied development economist that the one-shot risky investment game invokes loss aversion. He may therefore interpret the experimental results through the glasses 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:

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

Risk aversion is under EUT captured by the concavity of the utility function and a concave utility function is necessary to get interior solutions in the game for x with $0 < x^* < 30$. 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 sample population may be derived from the observed investment distribution based on the one-shot standard game.

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, this therefore leads to the same individual risk aversion parameter derived for each subject based on her/his optimal x^* allocation that would be identical across the three treatments. Using the one-shot game to measure risk aversion would then lead to no bias in the estimation of risk aversion:

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

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

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

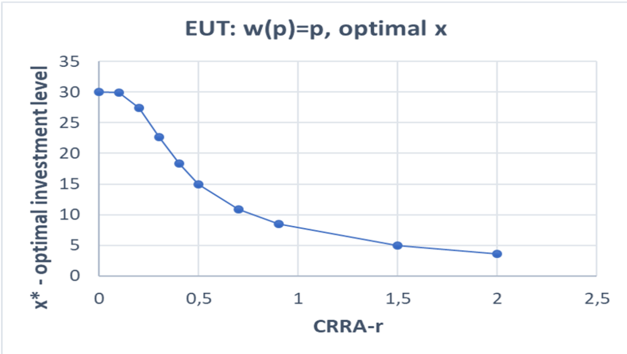

Fig. 1 EUT: CRRA-r and optimal investment in one-shot game

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.

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 decision point 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):

$$\max PT3(T1) = w^{+}(0.5)v(2x) - w^{-}(0.5)\lambda v(|x|) \quad (4)$$

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 in each of these treatments as they pull in opposite directions. They only allow us to verify whether they exist and are of a significant size. Holden and Tilahun (2021) demonstrated a substantial difference between the choices in T1 and T2. Building on the same data we introduced the new treatment T3.

Treatment T3 was designed to prevent endowment effects and was introduced as a number of binary choices without any allocation of endowment till after all the binary choices have been made. Based on the assumption of zero asset integration, the reference point is therefore zero for each binary choice. The elicitation approach is similar to how binary choices are introduced in MPLs or the strategy method frequently used to avoid income or endowment effects in experiments. Without the respondents knowing, the sequence of binary choices is designed to narrow in on the optimal mix of safe and risky amounts in the game when corner solutions are not preferred. We suggest that EUT may be used to measure risk aversion (utility curvature) for T3.

Significant treatment effects imply that EUT must be rejected. The dominance of interior choices point towards non-linear utility functions for relatively small amounts such as those used in this game, implying limited or no asset integration. Some recent findings in the experimental literature have found evidence of quite flat (close to linear) value functions (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). This may point towards a need to modify the theory or a need for an alternative theory. One option is to use a hybrid between EUT and PT. 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, sensitivity could be increasing in the loss domain.

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 trade in the games. The two variants of the theory use standard probability weighted aggregation (AEET1) or aggregate 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 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 (part of) the risky lottery opportunity when converting it to a safe amount. The sophisticated subjects maximize the following problem for T2:

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

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, separate and aggregate utilities of risky and safe amounts. 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:

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

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

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 AEET1 and AEET2 theories are better at describing the observed outcomes across treatments and subjects. EUT predicts no treatment effects and dominance of interior solutions and has already been rejected based on the finding of strong treatment effects. PT predicts significant treatment effects and dominance of corner solutions. AEET predicts significant treatment effects and high likelihood of interior solutions.

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. 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.

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

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.

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. Treatments T2 and T3 were randomized at group level for the sample of youth business groups and group members in the pilot district.

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 T1 and T2:

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

Subscript g represents group, subscript i represents individual, r_1 represents the estimated share invested in treatment T1, α_2 captures the combined endowment effects for T2 and T1 which pull in opposite directions. α_3 represents the endowment effect for treatment T1 with treatment T3 as the baseline. The difference α_2-α_3 is the estimated endowment effect for the risky lottery. 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.

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

Figure 1 shows the full sample investment distribution for all three treatments. The figure illustrates highly significant differences in distributions across the three treatments. Figure 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 solutions 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 an endowment effect associated with the allocation of the risky amount in treatment T2.

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

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

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

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

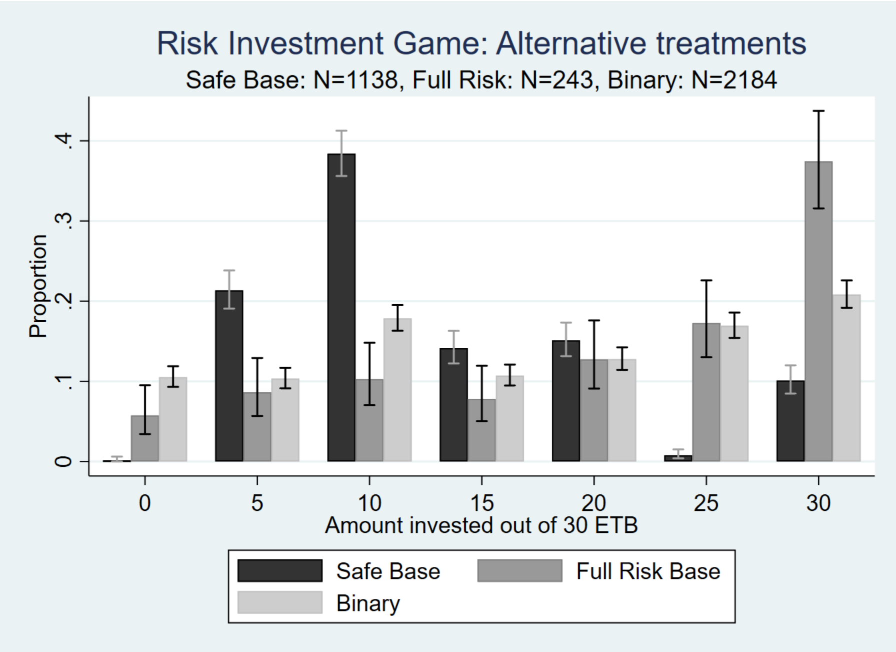

Fig. 2 Distribution of investments in Treatments 1, 2 and 3 (full sample)

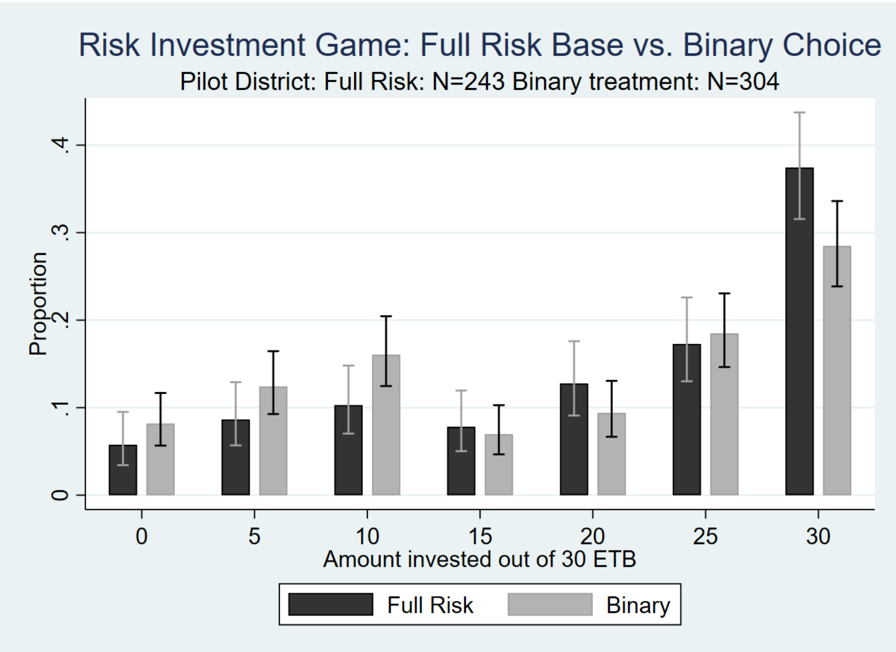

Fig. 3 Robustness check Treatment 2 (Full Risk) and Treatment 3 (Binary) in pilot district

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

Table 4 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)	0.206***	0.210***	0.244***	0.161***
	(0.029)	(0.029)	(0.042)	(0.039)
T3-Binary(#Female)	0.096***	0.101***	0.124***	0.111***
	(0.018)	(0.018)	(0.025)	(0.019)
Male		0.046***	0.070***	0.054***
		(0.012)	(0.020)	(0.018)
T2#Male			0.263***	
			(0.035)	
T3#Male			0.161***	
			(0.023)	
Age		-0.000	-0.000	-0.002**
		(0.001)	(0.001)	(0.001)
Birth rank		0.005**	0.006**	0.005
		(0.002)	(0.002)	(0.004)
Education (years)		0.006***	0.006***	0.002
		(0.001)	(0.001)	(0.002)
Constant	0.376***	0.315***	0.298***	0.396***
	(0.024)	(0.033)	(0.035)	(0.046)
Observations	3,564	3,564	3,564	1,487
Number of youth groups	308	308	308	305

All models with district FE and enumerator FE
T1 (Safe base) is baseline treatment (Constant)
Cluster-robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

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 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.

Table 5 Robustness checks for pilot district (Degua Tembien) sample

	(5)	(6)	(7)
VARIABLES	riskshare	riskshare	riskshare
Panel controls	Group RE	Group FE	Individual FE
T2-Full Risk	0.185***	0.197***	0.174**
	(0.040)	(0.047)	(0.067)
T3-Binary	0.105***	0.102**	0.114*
	(0.039)	(0.047)	(0.067)
Constant	0.429***	0.430***	0.410***
	(0.045)	(0.044)	(0.030)
Observations	822	822	822
R-squared		0.141	0.292
Number of groups	53	53	53
Number of individuals	593	593	593

All models with enumerator FE. Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

To further inspect the robustness of the results, the pilot district sample is used with alternative controls, see models (5) - (7) 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 treatment effects were robust to these alternative specifications. T2 and T3 give significantly larger average investment levels than T1 in all model specifications and the investment levels are 17.4-19.7 percentage points higher for T2 than for T1 and 10.2-11.4 percentage points higher for T3 than for T1. 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 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 in 2016 to T2 and T3 in 2019 cannot be explained by such shocks affecting the respondents and changing 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 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 measure 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 (PT) can explain and predict different behavior across treatments, it does not explain the dominance of interior solutions 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 0.7 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.

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 r for the three treatments with $\delta = 0.1$.

Figure 4 illustrates that for the AEET1 model interior solutions dominate for a wide range of CRRA-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.

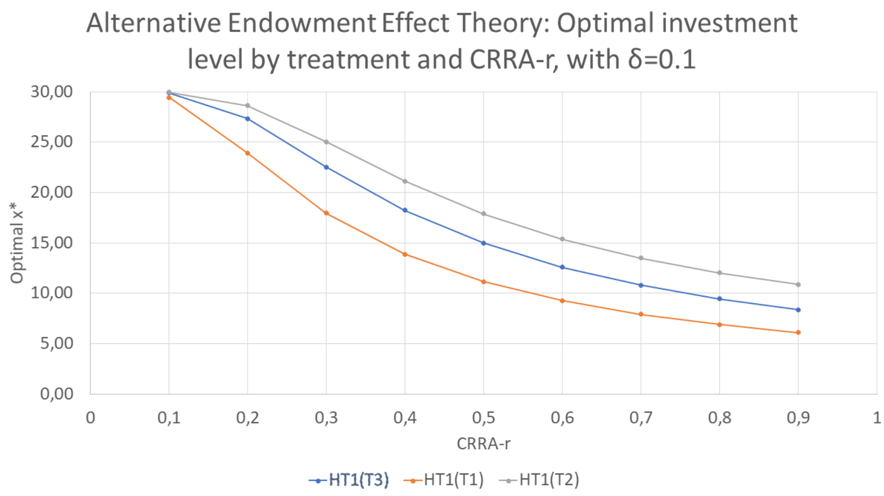

Fig. 4 Optimal investment (x*) by treatment and CRRA-r, with endowment effect

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 solutions 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 that are not feasible with the simple one-shot version.

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 and participants in the 42nd Annual Meeting of the Norwegian Association of Economists, NMBU, Ås, 6-7th January 2020.

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.

Cappelletti, D., Mittone, L., and Ploner, M. (2014). Are default contributions sticky? an experimental analysis of defaults in public goods provision. *Journal of Economic Behavior & Organization*, 108:331–342.

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.

Dhingra, N., Gorn, Z., Kener, A., and Dana, J. (2012). The default pull: An experimental demonstration of subtle default effects on preferences. *Judgment and Decision Making*, 7(1):69.

Epley, N. and Gilovich, T. (2001). Putting adjustment back in the anchoring and adjustment heuristic: Differential processing of self-generated and experimenter-provided anchors. *Psychological Science*, 12(5):391–396.

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.

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

Kahneman, D., Knetsch, J. L., and Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives*, 5(1):193–206.

Marzilli Ericson, K. M. and Fuster, A. (2011). Expectations as endowments: Evidence on reference-dependent preferences from exchange and valuation experiments. *The Quarterly Journal of Economics*, 126(4):1879–1907.

Plott, C. R. and Zeiler, K. (2007). Exchange asymmetries incorrectly interpreted as evidence of endowment effect theory and prospect theory? *American Economic Review*, 97(4):1449–1466.
Samuelson, W. and Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty*, 1(1):7–59.
Schmidt, U., Starmer, C., and Sugden, R. (2008). Third-generation prospect theory. *Journal of Risk and Uncertainty*, 36(3):203.
Simonson, I. and Drolet, A. (2004). Anchoring effects on consumers' willingness-to-pay and willingness-to-accept. *Journal of Consumer Research*, 31(3):681–690.
Sunstein, C. R. (2002). Switching the default rule. *NYUL Rev.*, 77:106.
Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, 1(1):39–60.

A Appendix 1.

A.1 Pilot district representativeness and individual controls

Figure 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.

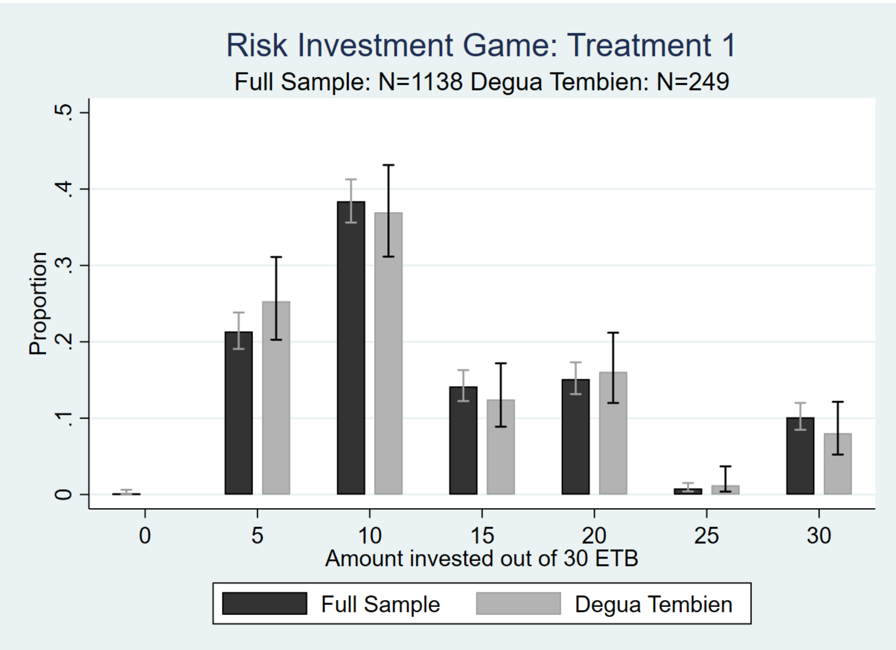

Fig. 5 Robustness check (Treatment 1): Pilot district vs full sample

Table A1 presents the individual control variables and tests for their difference across treatments. With Treatments 2 and 3 implemented three years later it is expected that the respondents on average will be about three years older in Treatments 2 and 3.

Table A1. Individual characteristics by treatment: t-tests

	T1	T2	T3	t-tests	t-tests	t-tests
	Safe Base	Full Risk	Binary	T1 vs T2	T1 vs. T3	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

A.2 Robustness checks and individual controls

We also ran separate regression models for this pilot district with the 2016 and 2019 samples jointly. 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. This allowed control for time-invariant group and individual unobservable characteristics. The following key alternative specifications were estimated to test the robustness of the results:

Our design confounds year with the baseline treatments and there is a risk that the youth have changed their behavior in the baseline treatment over this three year period. We scrutinized this in two 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 Treatment 1 versus Treatments 2 and 3. Appendix Table A2 assesses the correlation between individual controls and the risky investment in the pilot district);

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 (pilot district) and A4 (full sample)).

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 A2. Robustness check: Individual controls in Pilot district

	(1)	(2)	(3)
VARIABLES	riskshare	riskshare	riskshare
Panel controls	Group RE	Group FE	Individual FE
T2-Full Risk	0.188***	0.198***	0.174**
	(0.041)	(0.047)	(0.067)
T3-Binary	0.105***	0.100**	0.114*
	(0.040)	(0.048)	(0.067)
Male, dummy	0.041*	0.035	
	(0.025)	(0.027)	
Age	-0.001	-0.000	
	(0.001)	(0.002)	
Birth rank	0.009	0.008	
	(0.006)	(0.006)	
Education, years	0.006**	0.006*	
	(0.003)	(0.003)	
Constant	0.359***	0.355***	0.410***
	(0.065)	(0.068)	(0.030)
Observations	822	822	822
R-squared		0.149	0.292
Number of groups	53	53	53
Number of individuals			593

All models with enumerator FE. Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A3. Robustness check: Recent shock effect in pilot district

	(1)	(2)	(3)
VARIABLES	riskshare	riskshare	riskshare
Panel controls	Group RE	Group FE	Individual FE
T2-Full Risk	0.182***	0.193***	0.181***
	(0.041)	(0.047)	(0.070)
T3-Binary	0.102***	0.101**	0.117*
	(0.039)	(0.047)	(0.067)
2018-19 Shock dummy	0.032	0.031	-0.043
	(0.043)	(0.045)	(0.097)
Constant	0.426***	0.427***	0.411***
	(0.045)	(0.045)	(0.031)
Observations	822	822	822
R-squared		0.141	0.292
Number of groups	53	53	53
Number of individuals			593

All models with enumerator FE. Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A4. Robustness check: Recent shock variable included

	(1)	(2)	(3)
VARIABLES	riskshare	riskshare	riskshare
Sample	Full	Full	Same enumerators
T2-Full risk	0.207***	0.211***	0.163***
	(0.029)	(0.029)	(0.040)
T3-Binary	0.096***	0.101***	0.112***
	(0.018)	(0.018)	(0.019)
2018-19 Shock dummy	-0.008	-0.003	-0.016
	(0.022)	(0.022)	(0.037)
Male dummy		0.046***	0.054***
		(0.012)	(0.018)
Age		-0.000	-0.002**
		(0.001)	(0.001)
Birth rank		0.006**	0.005
		(0.002)	(0.004)
Education (years)		0.006***	0.002
		(0.001)	(0.002)
Constant	0.377***	0.315***	0.396***
	(0.024)	(0.033)	(0.046)
Observations	3,564	3,564	1,487
Number of groups	308	308	305

Model (3) Same enumerators in 206 and 2019
All models with enumerator FE and group RE
Standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1