

Measurement errors of risk aversion and how to correct them^{*}

Christoph Engel[†] Oliver Kirchkamp[‡]

17th March 2018

We provide an example for measurement errors which is common but also often neglected in economic experiments: In one task participants' attitudes are measured, in another task participants' behaviour is related to this measurement. How should we deal with imperfect measurements of these attitudes?

To correct for measurement errors one must first assess the magnitude of the error. To do this one needs several measurements. Since repeated measurements are deemed expensive and complicated, researchers tend to ignore the problem.

Here we show that (a) the problem exists and (b) if risk attitudes are elicited with the help of a multiple price list task, then a simple solution exists. Seemingly inconsistent answers in the task can be exploited to assess and to correct for the measurement error.

We illustrate the problem with an experiment which studies the relation between risk aversion and punishment behaviour.

Keywords: Measurement error, risk, lab experiment, public good, Bayesian inference.

JEL: C91, D43, L41

1. Introduction

When we run laboratory experiments and when we try to structure the results of these experiments, we sometimes combine two parts of an experiment. In one part of the experiment we measure individual attitudes. These measurements are used to explain behaviour in another part of the experiment. Clearly, one can not assume that these measurements are free

^{*}Helpful comments by Ioanna Grypari and André Schmelzer on an earlier version are gratefully acknowledged. Data and methods used in the paper can be found at <http://www.kirchkamp.de/research/errorsInVar.html>

[†]MPI for Research on Collective Goods, Kurt-Schumacher-Str. 10, 53113 Bonn, engel@coll.mpg.de, +49 228 91416-0.

[‡]University Jena, School of Economics, Carl-Zeiss-Str. 3, 07737 Jena, oliver@kirchkamp.de, +49 3641 943240.

of any errors. Traits and attitudes are unlikely to be stable across situations (Ross, Nisbett and Gladwell, 2011). Attitudes will often only be imperfectly observed in post-experimental tests.

For the econometrician the problem of an explanatory variable that is measured only with some error is well known as one of “errors in variables”. If we want to estimate $Y = \beta_0 + \beta_1 \xi + u$, but we can observe ξ only with an error, e.g. we observe $X \sim \mathcal{N}(\xi, \sigma_X)$, then estimating $Y = \beta_0 + \beta_1 X + u$ with standard OLS provides a biased and inconsistent estimator for β_1 . Already Adcock (1877) mentions the problem of measurement errors. Since then many authors have contributed to the discussion (see Gillard, 2010, for an overview).

In the above problem our estimate for β_1 is biased towards zero. The measurement error makes it harder to find an effect if an effect exists. If, nevertheless, we find something, then the true effect must be even stronger. In this context measurement errors might be considered harmless.

The situation changes when further variables come into play. Let us assume we have two explanatory variables, i.e. we estimate $Y = \beta_0 + \beta_1 \xi_1 + \beta_2 \xi_2 + u$. Let us further assume that ξ_1 and ξ_2 are positively correlated and ξ_1 can only be observed with an error, i.e. we only observe $X_1 \sim \mathcal{N}(\xi_1, \sigma_{X_1})$. Then estimating $Y = \beta_0 + \beta_1 X_1 + \beta_2 \xi_2 + u$ with standard OLS underestimates β_1 and overestimates β_2 . Neglecting the error in measuring ξ_1 might wrongly lead us to believe that ξ_2 would matter, even when in reality ξ_2 is irrelevant.

Hey and Orme (1994) points out the importance of error in understanding the relation between preferences and economic decisions. The relevance of measurement errors for risk preferences has been highlighted in a number of studies (Harrison et al. 2005; Loomes 2005; Sahm 2012, to name only a few). Kimball, Sahm and Shapiro (2008) use an error in variables model to interpret survey data on risk tolerance. Beauchamp, Cesarini and Johannesson (2017) employ a latent variable model to assess the predictive power of risk-attitudes. Gilen, Snowberg and Yariv (2015) replicate three influential experimental studies (Niederle and Vesterlund 2007, Friedman et al. 2014, Halevy 2007), pointing out that in all these studies accounting for measurement error changes the interpretation of the results.

Still, the number of economic experiments which take into account measurement errors is small. One reason might be that to correct for measurement errors one has to assess the magnitude of the error. This assessment requires multiple measurements. Multiple measurements are more expensive and more complicated than single measurements.

Here we propose a simpler approach: Instead of taking the measurement several times we exploit inconsistencies in behaviour to estimate the precision of our measurement. More specifically, we exploit that a very common task to measure risk aversion by Holt and Laury (2002) is based on a multiple price list, i.e. a task which by construction requires multiple measurements.

To demonstrate our approach we consider the following example: In one part of an experiment we measure participants’ attitudes towards risk with the help of a Holt and Laury (2002) task. Inconsistent answers, which are often considered to be a burden, turn out to be a blessing as they allow us to assess the magnitude of the measurement error. We then use the risk measure as an explanatory variable in a second task, here behaviour in a public good game with punishment. We show how a joint estimation of both decision processes

has two advantages: First, one uses the data from all participants, thus avoiding a selection bias.¹ Second, we can estimate, separately for each participant, the precision of the measure for her risk attitude. This allows us to take into account the measurement error. We can show that results change substantially if one respects the measurement error in the Holt and Laury task.

The remainder of the paper is organized as follows: Section 2 introduces the design of the example experiment from which the data are taken and that we use to illustrate our methodological point. Section 3 discusses alternative methods for dealing with inconsistency in the measurement of risk attitudes. Section 4 uses simulations to assess the size of the bias due to the measurement error in a more general context. Section 5 concludes.

2. Design of the Example Experiment

In our example we revisit the data from a four-person repeated public good game with punishment (Engel, 2014). The experiment was conducted in the Cologne Laboratory for Economic Research in 2012. The experiment was implemented in zTree (Fischbacher, 2007). Participants were invited using the software ORSEE (Greiner, 2004). Of 90 participants 80 were students of various majors with a mean age 25.4. 44% were female. Participants on average earned 15.11€ (19.82\$ on the days of the experiment), 14.80€ for players in the public good game, and 16.38€ for players with the power to punish (authorities). The experiment had 3 sessions of 30 participants (6 groups of 4 active participants; 6 passive authorities).

The aim of the experiment was to study the relation between attitudes to risk and the reaction to punishment in a public good game. Fehr and Gächter (2000) show that if participants in a public good game have the possibility to punish each other, contributions in the public good game stabilize at a high level.

In the experiment four (active) participants $i \in \{1, \dots, 4\}$ in group k contribute c_{ikt} in round t to a linear public good. A fifth participant α (an authority) has the power to impose a punishment η_{ikt} on each active participant. Profits π_{ikt} of the active participants and π_{akt} of the authority are given by (1) and (2):

$$\text{Profit of active participant } i: \quad \pi_{ikt} = 20 - c_{ikt} + .4 \sum_i c_{ikt} - 3\eta_{ikt} \quad (1)$$

$$\text{Profit of authority } \alpha: \quad \pi_{akt} = 25 + 20 - \sum_i \eta_{ikt} \quad (2)$$

The game is repeated for 11 periods. Participants are rematched every period to matching groups of size 10.

After the main experiment, Engel (2014) administers (among several other tests) a test for risk aversion, following Holt and Laury (2002). Holt and Laury design a task where participants choose between a (safe) lottery with a small spread, $p \cdot 2\$ + (1 - p) \cdot 1.6\$$, and a

¹Otherwise one does not estimate the effect of risk aversion on punishing behavior in the population, but the effect of risk aversion on the punishing behavior of only those individuals whose reactions to risky choices are highly consistent.

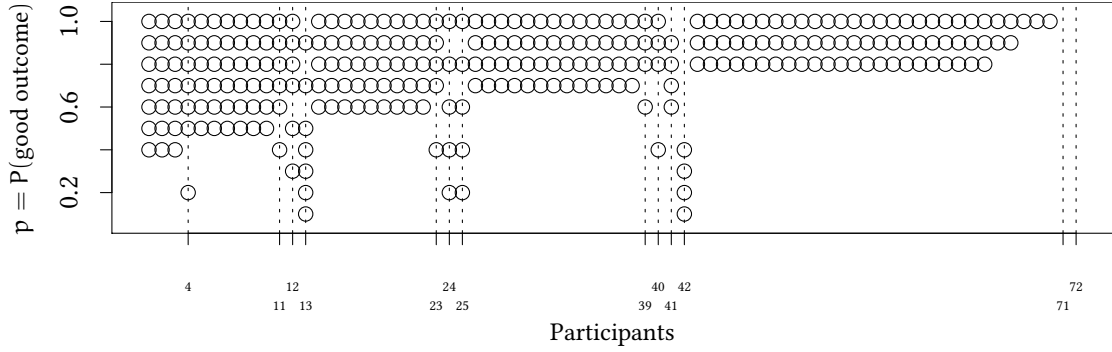


Figure 1: Choices in the risk task

The panel shows choices for each participant: \circ if the participant choses the lottery with the larger spread (risky) and nothing if the participant chose the smaller spread (safe). Vertical reference lines denote participants with inconsistent choices (see (4)). Participants are ordered by their risk attitudes, with the more risk seeking participants at the left.

(risky) lottery with a large spread, $p \cdot 3.85\$ + (1 - p) \cdot .1\$$, where the probability of the good outcome is $p \in \{.1, .2, .3, \dots, 1\}$.

In this experiment, Engel (2014) is interested in the impact of punishment on future cooperation. Punishment in the most recent period is the most vivid experience and should have the strongest effect. The more they are averse to risk, the stronger this signal should guide choices in the subsequent period: risk averse participants loose more utility when punished again. Engel (2014) investigates the following hypothesis:

Hypothesis 1 *The more a participant is risk averse, the more she increases her contributions to a linear public good after having been punished in the previous period.*

3. How to Deal with an Inconsistent Measure for Risk Attitudes?

3.1. Measuring risk aversion

To test Hypothesis 1, we need for each active participant a measure of her risk aversion. Choices in the risk task for the 72 participants holding the active role are shown in Figure 1. Choices where a participant chose the lottery with the larger spread (risky) are denoted with a \circ , choices where the participant chose the lottery with the smaller spread (safe) are left blank.

If we assume that preferences for money follow e.g. CRRA, i.e. $u(z) = z^{1-r}$, then the critical value of p^c where participants are indifferent between the more safe and the more risky choice is a monotonic function of their relative risk aversion r . We can then either describe participants by r or by their critical value of p^c . In the following we will use p_{ik}^c to describe preferences of individual i in group k .

In Figure 1. we have ordered participants from the most risk loving on the left to the most risk averse on the right. If we could describe decision makers by a single switching point p_{ik}^c then the following holds:

$$\text{choice}_{ik}(p) = \begin{cases} \text{risky} & \text{if } p > p_{ik}^c \\ \text{either safe or risky} & \text{if } p = p_{ik}^c \\ \text{safe} & \text{if } p < p_{ik}^c \end{cases} \quad (3)$$

We call a participant i in group k *consistent* if there is a p_{ik}^c which rationalises all choices, i.e.

$$\max\{p | \text{choice}_{ik}(p) = \text{safe}\} \leq p_{ik}^c \leq \min\{p | \text{choice}_{ik}(p) = \text{risky}\}. \quad (4)$$

We call a participant *inconsistent* if (4) does not hold, i.e. not all their choices can be rationalised with Equation (3). In Figure 1 vertical reference lines denote participants with inconsistent choices. The choices of 18% of the participants in Engel (2014)'s experiment are inconsistent.

One way to formalise consistent and inconsistent choices in the Holt and Laury task is the logistic model. The probability of a risky choice of individual ik in lottery p could be written as follows:

$$P(\text{risky}_{ik}|p) = \mathcal{L}((p - p_{ik}^c) \cdot \sqrt{\tau_{ik}}) \text{ with } p \in \{.1, .2, \dots, 1\} \quad (5)$$

Here \mathcal{L} is the logistic function, p describes the probability of the good outcome in the Holt and Laury task, p_{ik}^c is the critical value where participant ik is just indifferent between the two choices, and τ_{ik} is the precision of the observation.

While $\tau_{ik} = \infty$ would guarantee us only consistent choices, we will show below that a value for τ_{ik} between about 2 and 60 describes behaviour better. The smaller τ_{ik} , the more frequent are inconsistent choices.

One possible reaction to inconsistent choices is using an alternative test that forces consistency. Eckel and Grossman (2008) directly ask participants for the switching point. Depending on the research question, this may be satisfactory. However, by enforcing consistent choices for p_{ik}^c we lose information about the decision maker's precision τ_{ik} of that choice. Below we will argue that information about this precision may be useful.

Before we do this, let us come to the contributions in the public good game.

3.2. Contributions to the public good

To test Hypothesis 1 we have to explain changes in the contribution to the public good Δc_{ikt} as a function of previous punishment η and risk aversion p_{ik}^c . We eventually want to estimate the following model:

$$\Delta c_{ikt} = \beta_0 + \beta_\eta \eta_{ik,t-1} + \beta_p p_{ik}^c + \beta_{\eta \times p} \eta_{ik,t-1} \cdot p_{ik}^c + \nu_k + \nu'_{ik} + \epsilon_{ikt} \quad (6)$$

Δc_{ikt} is the change of contribution to the public good of individual i from matching group k at time t . $\eta_{ik,t-1}$ is the punishment received by individual i from group k at time $t-1$, i.e. the punishment received in the previous period. p_{ik}^c is a measure for risk aversion of individual

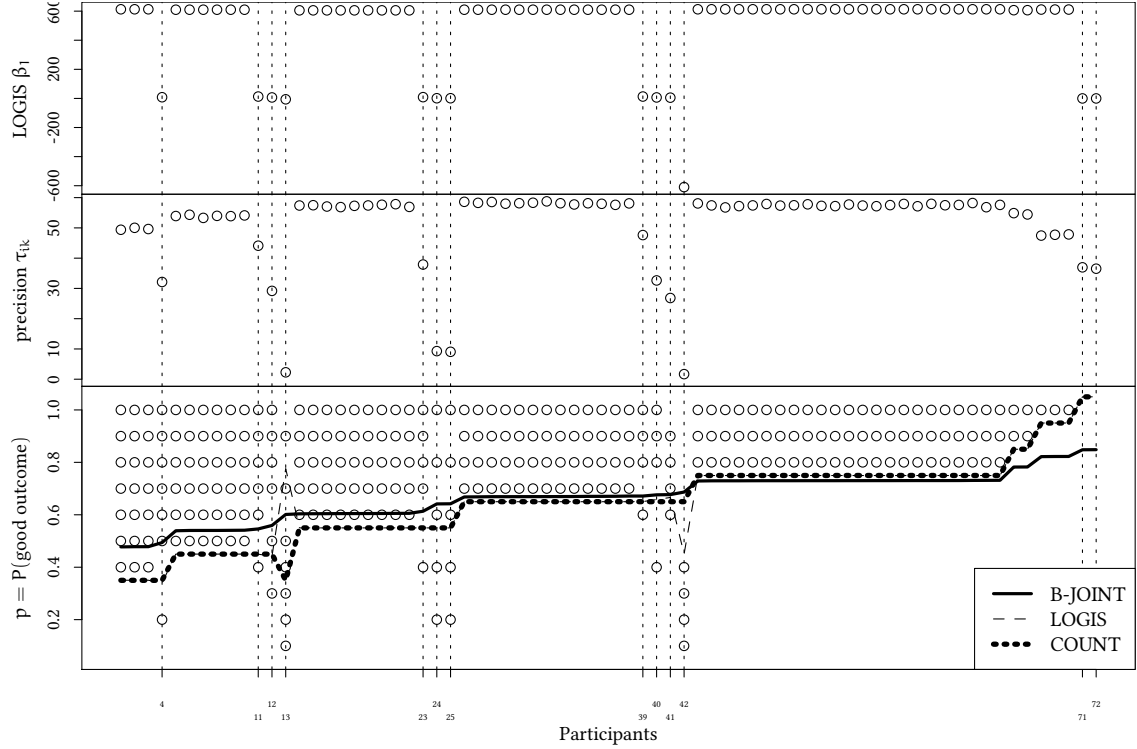


Figure 2: Choices, switching points p_{ik}^c and precision of choice τ_{ik}

The bottom panel shows for each participant the actual choices: \circ if the participant chose the more risky lottery. Participants are ordered by their median switching points p_{ik}^c as estimated from the B-JOINT model. The solid line denotes the median estimated switching points p_{ik}^c from B-JOINT. The dashed line shows the estimated switching points from LOGIS. Vertical reference lines denote participants with inconsistent choices, i.e. with more than one switching point. The panel in the middle shows the estimated values of the participant's precision, τ , from B-JOINT. The top panel shows the estimated value of β_1 from LOGIS.

i from group k . ν_k is a random effect for group k . ν'_{ik} is a random effect for individual i from group k . ϵ_{ikt} is the residual. In line with Hypothesis 1 we expect the interaction term $\beta_{\eta \times p}$ to be positive.

In the example study, the aim is to explain reactions to punishment as a function of the attitude towards risk. The latter is described as a switching point p_{ik}^c in the Holt and Laury task. In Section 3.3 we briefly discuss four alternative approaches to deal with inconsistent choices. All four approaches can be used to estimate Equation (6), but all assume that p_{ik}^c could be measured with infinite precision. As a result all four approaches neglect the measurement error. In Section 3.4 we estimate the decision process determining p_{ik}^c jointly with Equation (6). This approach offers a solution for the measurement error.

3.3. No correction for the measurement error

3.3.1. Drop inconsistent observations (DROP)

This procedure removes from our sample the 18% of the participants which are inconsistent according to (4). In Figure 2 these are the participants which are crossed out by a vertical dashed line. For the remaining 82% of our participants we define the switching point as follows:

$$\hat{p}_{it}^{c,D} = \frac{\max\{p|\text{choice}_{ik}(p) = \text{safe}\} + \min\{p|\text{choice}_{ik}(p) = \text{risky}\}}{2} \quad (7)$$

Figure 2 suggests that inconsistent behaviour could be more likely with risk seeking participants. The DROP procedure might, hence, selectively remove risk seeking participants from the sample. It also does not tell us anything about the precision of p_{ik}^c , i.e. it does not help us to address the measurement error.

3.3.2. Counting the number of safe choices (COUNT)

Holt and Laury (2002) propose to replace the switching point for inconsistent participants by simply counting the number of safer choices. To ease the comparison with the other measures we use the following linear transformation:

$$\hat{p}_{it}^{c,C} = \frac{1}{20} + \frac{1}{10} \sum_p [\text{choice}_{ik}(p) = \text{safe}] \quad (8)$$

Figure 2 shows the resulting estimates of risk preferences as a thick dotted line. This procedure addresses the selection bias but not the measurement error.

3.3.3. A logistic regression to estimate switching points (LOGIS)

In Equation (5) we have used the logistic model to describe choices in the risk task. We can rephrase this model as follows:

$$P(\text{risky}_{ik}|p) = \mathcal{L}(\beta_{0,ik} + \beta_{1,ik}p) \text{ where } p \in \{.1, .2, \dots, 1\} \quad (9)$$

The value of p where the $P(\text{risky}_{ik}|p) = 1/2$, i.e. where individual i in group k chooses the more risky and the safer lottery with equal probabilities, is our estimated switching point $\hat{p}_{ik}^{c,L}$. It is given by

$$\hat{p}_{ik}^{c,L} = -\hat{\beta}_{0,ik}/\hat{\beta}_{1,ik} \quad (10)$$

The dashed line in the bottom part Figure 2 shows for each individual the critical value $\hat{p}_{ik}^{c,L}$ obtained with this method.² As Figure 2 demonstrates, the results obtained with LOGIS are

²Note that LOGIS (the same way as the Bayesian methods) easily handles “inconsistent” participants. Figure 1 shows that we have 13 such participants in the dataset. We have no participants who, independent of p , always choose the risky lottery. These participants would correspond to $\hat{p}_{ik}^{c,L} < 0$. We have two participants who always choose the safe lottery. They correspond to $\hat{p}_{ik}^{c,L} > 1$.

similar to COUNT, except for participants 13 and 42.³ The top part of the same figure shows for each individual the coefficient $\hat{\beta}_{1,ik}$. When this coefficient is large, then $P(\text{risky}_{ik}|p)$ is either close to 1 or close to 0 for most values of p . A large coefficient is, hence, a measure of consistency. When we use maximum likelihood to estimate Equation (9) we should expect that for consistent choices $\hat{\beta}_{1,ik} \rightarrow \infty$. Since numerical precision is limited we find for consistent choices in our estimation $605 \leq |\hat{\beta}_{1,ik}| \leq 613$ which is clearly smaller than $+\infty$, but already sufficiently large to make sure that the actual choices are made almost with certainty.⁴ Still, we should keep in mind that it is only numerical imprecision which yields finite values where we should see a $+\infty$.

Looking at Figure 2 again we see two (related) problems:

1. For the 18% inconsistent choices we have $\hat{\beta}_{1,ik} \leq 13$. These choices are clearly more noisy than the 82% consistent choices with $\hat{\beta}_{1,ik} \geq 605$ but it is not obvious how to exploit this difference in precision in our estimate of Equation (6).
2. The estimation of Equation (9) yields for two participants (13 and 42) negative values for $\hat{\beta}_1$ (-6.1 and -610). These participants choose the safer lottery more frequently when the probability of the good outcome is larger. The LOGIS model does not tell us how to interpret such, seemingly perverse, choices.

We will argue below that these 18% inconsistent participants can serve two purposes. First, although their observations are noisy, dropping them would lead to a selection bias. Second, and more importantly, the noise of these observations allows us to take into account the measurement error. If 18% of our participants clearly violate consistency, we should, perhaps, not expect that we can measure the remaining 82% with infinite precision. The inconsistent 18% will allow us to better assess the precision of the remaining 82% consistent observations.

3.3.4. Estimation results for DROP, COUNT and LOGIS

Table 1 shows the estimation results for Equation (6) for different ways to deal with inconsistent observations. We see that, regardless which method we use here, the differences are not very large. We find β_η between 1.19 and 1.36, β_p is never significant and between 0.339 and 1.13, and $\hat{\beta}_{\eta \times p}$ somewhere between -1.17 and -0.876.

Irrespective of the estimation procedure, a perfectly risk loving subject ($p^c = 0$) increases her contributions by more than 1 unit in response to any unit of punishment she has received in the previous period. The more the participant is risk averse, the less intense her reaction. Yet even a perfectly risk averse participant ($p^c = 1$) still exhibits a small increase

³Since the logistic model is not fully identified it is only a convenient artefact of the numerical implementation to find a unique answer to the question for the optimal switching point. If a participant has chosen the safer lottery for all choices $p \leq .6$ and the more risky lottery for all choices $p \geq .7$, the logistic model will estimate a switching point just in the middle between .6 and .7 at almost exactly .65.

⁴If a participant is just indifferent at p^c , i.e. $\beta_0 + \beta_1 p^c = 0$, then the next actual choice in the experiment is made for $p = p^c + 1/20$ and $p = p^c - 1/20$. The probability of a safe or risky choice there is, hence, $\mathcal{L}(\beta_{1,ik}/20)$ and $\mathcal{L}(-\beta_{1,ik}/20)$. For $\beta_{1,ik} = 605$ we have $\mathcal{L}(605/20) \approx 1 - 7.29 \times 10^{-14}$, $\mathcal{L}(-605/20) \approx 7.29 \times 10^{-14}$.

| | DROP | | | COUNT | | | LOGIS | | |
|------------------|------------|----------|--------------|------------|----------|--------------|------------|----------|--------------|
| | β | 2.5% | 97.5% | β | 2.5% | 97.5% | β | 2.5% | 97.5% |
| 0 | -1.267 | -2.776 | 0.171 | -0.809 | -2.025 | 0.423 | -0.732 | -1.962 | 0.496 |
| η | 1.356 | 0.663 | 2.050 | 1.208 | 0.699 | 1.702 | 1.189 | 0.673 | 1.692 |
| p | 1.130 | -0.944 | 3.291 | 0.457 | -1.322 | 2.190 | 0.339 | -1.427 | 2.085 |
| $\eta \times p$ | -1.172 | -2.198 | -0.145 | -0.909 | -1.632 | -0.166 | -0.876 | -1.608 | -0.126 |
| | σ^2 | σ | $1/\sigma^2$ | σ^2 | σ | $1/\sigma^2$ | σ^2 | σ | $1/\sigma^2$ |
| ν'_{ik} | 0.000 | 0.000 | Inf | 0.000 | 0.000 | Inf | 0.000 | 0.000 | Inf |
| ν_k | 0.287 | 0.536 | 3.483 | 0.375 | 0.612 | 2.666 | 0.379 | 0.616 | 2.638 |
| ϵ_{ikt} | 10.381 | 3.222 | 0.096 | 10.296 | 3.209 | 0.097 | 10.301 | 3.210 | 0.097 |

Table 1: ME estimates of Equation (6).

of contributions in reaction to punishment ($0.184 \leq \beta_\eta + \beta_{\eta \times p} \leq 0.313$ depending on the model).

3.4. Accounting for the measurement error – joint estimation of (5) and (6)

The previous three approaches treat the estimation p_{ik}^c from Equation (5) and the estimation of the impact of p_{ik}^c on Δc_{ikt} from Equation (6) as two unrelated problems. Here we suggest that much can be gained if both problems are estimated together. We will use a Bayesian approach. We do not want to enter a discussion on the comparative merits of the Bayesian versus the frequentist framework (Bayarri and Berger, 2004, or Kass, 2011 may provide a starting point for a discussion).

In the frequentist world the problem that we outline below could be described as a generalised multilevel structural equation (Rabe-Hesketh, Skrondal and Pickles, 2004). Neyman and Scott (1948) and Solari (1969) have pointed out deficiencies in the maximum likelihood approach to estimate models with measurement errors. Bayesian estimation has been shown to work well in the context of measurement errors for a long time and for a wide range of situations.⁵ During the last decades Markov chain Monte Carlo methods have become a powerful and accessible tool for Bayesian inference. Thus, the Bayesian approach lends itself particularly well to estimate models with measurement errors.

Here we employ the Bayesian approach, in particular since it facilitates a transparent description of the two processes we want to estimate jointly. To ease the comparison with the frequentist framework we base our estimations on vague priors.⁶ We will present the joint model in Section 3.4.1. We will then demonstrate in Section 3.4.2 that, as long as the frequentist and the Bayesian approach estimate similar models, the results are (as they should) almost indistinguishable.⁷ We will then show in Section 3.4.3 that, once the measurement

⁵Arminger and Muthén (1998), Dellaportas and Stephens (1995), Florens, Mouchart and Richard (1974) and Polasek and Krause (1993).

⁶For a frequentist analysis the Stata package `gllamm` or the R package `lavaan` might be useful.

⁷We use JAGS 4.0.0. to estimate Bayesian models. Estimates are based on four chains with each 1000 samples

error is taken into account, results change substantially.

3.4.1. The joint model

Likelihoods: The likelihood of the Holt and Laury task is given by Equation (5). We re-write Equation (6) to obtain the likelihood for the public good task as follows:

$$\Delta c_{ikt} \sim \mathcal{N}(\beta_0 + \beta_\eta \eta_{ik,t-1} + \beta_p p_{ik}^c + \beta_{\eta \times p} \eta_{ik,t-1} \cdot p_{ik}^c + v_k + v'_{ik}, 1/\sqrt{\tau_\epsilon}) \quad (11)$$

Priors: We use the following (vague) priors:⁸

For the coefficients from Equation (11):

$$\beta_l \sim \mathcal{N}(0, 100) \text{ with } l \in \{0, \eta, p, \eta \times p\} \quad (12)$$

The group specific random effect in Equation (11):

$$v_k \sim \mathcal{N}(0, 1/\sqrt{\tau_v}); \text{ with } \tau_v \sim \Gamma(m_v^2/d_v^2, m_v/d_v^2); m_v \sim \Gamma(1, 1); d_v \sim \Gamma(1, 1) \quad (13)$$

The individual specific random effect in Equation (11):

$$v'_{ik} \sim \mathcal{N}(0, 1/\sqrt{\tau_{v'}}); \text{ with } \tau_{v'} \sim \Gamma(m_{v'}^2/d_{v'}^2, m_{v'}/d_{v'}^2); \\ m_{v'} \sim \Gamma(1, 1); d_{v'} \sim \Gamma(1, 1) \quad (14)$$

For the switching point from the risk task, Equation (5):

$$p_{ik}^c \sim \mathcal{B}(\alpha_c, \beta_c) \text{ with } \alpha_c \sim \Gamma(2, 1/2); \beta_c \sim \Gamma(2, 1/2) \quad (15)$$

For the precision of the switching point from Equation (5):

$$\tau_{ik} \sim \Gamma(m^2/d^2, m/d^2); \text{ with } m \sim \Gamma(1, 1); d \sim \Gamma(10, 0.1) \quad (16)$$

The precision in Equation (11):

$$\tau_\epsilon \sim \Gamma(m_\epsilon^2/d_\epsilon^2, m_\epsilon/d_\epsilon^2); \text{ with } m_\epsilon \sim \Gamma(1, 1); d_\epsilon \sim \Gamma(1, 1) \quad (17)$$

While this notation might look a bit intricate for the novice, priors and likelihoods can in a straightforward way be represented in JAGS or BUGS.⁹

3.4.2. Replicating LOGIS (B-LOGIS)

Before we come to the results of the joint estimation, let us use the Bayesian framework to replicate the result of the mixed effect estimation of Equation (6). As above we would treat both problems as unrelated. We would first estimate p_{ik}^c for each participant (using Equation (5), ignoring the public good game given by (11)). We would then, as if it was an independent problem, estimate Equation (11), ignoring the risk assessment (5). For both steps we use priors given by (12)-(17). Since the two problems are treated as unrelated, this procedure,

for adaptation, 4000 samples for burnin, and then, for each of the four chains, 100000 actual samples per chain. To estimate the mixed effects model we use `lme4 1.1-14`. Frequentist confidence intervals are based on normal bootstraps with 1000 samples.

⁸We use $\mathcal{N}(\mu, \sigma)$ for the normal distribution, $\Gamma(\alpha, \beta)$ for the Gamma distribution and $\mathcal{B}(\alpha, \beta)$ for the Beta distribution. The second argument of $\mathcal{N}(\mu, \sigma)$ is the standard deviation. $\tau = 1/\sigma^2$ is the precision. The first argument of $\Gamma(\alpha, \beta)$ is shape α , the second is rate β .

⁹See <http://www.kirchkamp.de/research/errorsInVar.html>

| B-LOGIS | | | | | |
|-----------------|--------|--------|--------|-------|--------|
| | Mean | 2.5% | 97.5% | SSeff | psrf |
| 0 | -0.757 | -2.103 | 0.572 | 4000 | 0.9998 |
| η | 1.212 | 0.697 | 1.726 | 3824 | 1.0002 |
| p | 0.356 | -1.369 | 2.388 | 3818 | 1.0000 |
| $\eta \times p$ | -0.890 | -1.640 | -0.158 | 3943 | 1.0005 |
| τ_v | 7.966 | 1.372 | 21.299 | 4118 | 0.9998 |
| τ'_v | 2.234 | 0.376 | 4.704 | 4000 | 1.0010 |
| τ_ϵ | 0.097 | 0.088 | 0.108 | 4000 | 0.9999 |

Table 2: Estimating Equations and (5) and (11) independently in the Bayesian Framework. No correction is made for the measurement error. Results are, as they should be, quite similar to the LOGIS or the COUNT model. We use 4 chains with 1000 samples each.

| B-JOINT | | | | | | B-JOINT-DROP | | | | | |
|-----------------|--------|--------|--------|-------|--------|-----------------|--------|--------|--------|-------|--------|
| | Mean | 2.5% | 97.5% | SSeff | psrf | | Mean | 2.5% | 97.5% | SSeff | psrf |
| 0 | -1.765 | -2.921 | -0.513 | 3916 | 1.0002 | 0 | -2.479 | -4.044 | -1.052 | 3843 | 1.0007 |
| η | 3.248 | 2.223 | 4.325 | 3891 | 0.9999 | η | 4.248 | 2.982 | 5.468 | 3913 | 1.0010 |
| p | 1.745 | 0.253 | 3.445 | 4000 | 1.0001 | p | 2.785 | 0.725 | 4.857 | 3905 | 1.0013 |
| $\eta \times p$ | -3.640 | -5.099 | -2.114 | 4334 | 0.9998 | $\eta \times p$ | -5.156 | -6.883 | -3.256 | 4105 | 1.0010 |
| τ_v | 6.894 | 1.486 | 16.797 | 4411 | 1.0022 | τ_v | 5.433 | 1.024 | 12.669 | 4041 | 1.0002 |
| τ'_v | 2.330 | 0.456 | 5.053 | 4265 | 1.0004 | τ'_v | 2.551 | 0.402 | 5.439 | 4257 | 1.0004 |
| τ_ϵ | 0.109 | 0.097 | 0.122 | 4000 | 0.9999 | τ_ϵ | 0.114 | 0.100 | 0.128 | 4000 | 1.0004 |

Table 3: Joint estimation of Equations (5) and (11).

The joint estimation corrects for the measurement error. The B-JOINT model uses all data (left table). B-JOINT-DROP uses only consistent participants (right table). We sample from 4 chains with 1000 samples each.

which we call B-LOGIS, ignores the measurement errors from the estimation of (5) when estimating (11). Estimation results are shown in Table 2. Here the value for $\beta_{\eta \times p}$ is -0.89, i.e. similar to the corresponding estimate of the mixed effects model based on the LOGIS estimate of p_{ik}^c ($\beta_{\eta \times p} = -0.876$). Also the value for β_η is with 1.21 similar to the LOGIS estimate ($\beta_{\eta \times p} = 1.19$). Finally, also the value for β_p is with 0.356 similar to the LOGIS estimate ($\beta_{\eta \times p} = 0.339$). All this should be reassuring: If the Bayesian and the frequentist framework have to solve similar problems, then both get very similar results.

3.4.3. B-JOINT

Next we jointly estimate Equations (5) and (11). This approach automatically weighs the individual estimates of the risk attitude by their precision and, thus, takes into account the measurement error. Priors are as given by Equations (12)-(17).

Equation (11): Estimation results (for the entire data set with 72 observations) are shown in the left part of Table 3. Figure 3 shows the highest posterior density (HPD) and confidence intervals for our estimates. The Figure illustrates the bias when not correcting for errors

in the measurement of risk. The joint estimate of B-JOINT finds (in absolute terms) a substantially larger effect size for $\beta_{\eta \times p}$ (-3.64) than the estimates we got from the models where we did not control for the measurement error (between -1.17 and -0.876). In other words: Correcting for the measurement error (and thereby weighting the individual measure of risk attitude with its precision) changes the effect size by 210%.

Equation (5): Figure 2 shows the predicted switching points p_{ik}^c as a solid line. The B-JOINT estimate for p_{ik}^c follows the estimates based on COUNT or LOGIS, in particular for the central values of p^c . For participants where LOGIS and COUNT estimate more extreme values of p^c , B-JOINT takes a more conservative approach. E.g. the extreme risk aversion of the rightmost participants in Figure 2 is not really in line with the distribution of the remaining values of p_{ik}^c . B-JOINT estimates, hence, a smaller precision τ_{ik} , and, accordingly, adjusts the value of p_{ik}^c more towards the centre of the distribution.

For individuals 13 and 42 (those, who choose the safer lottery more frequently when the probability of the good outcome was larger) LOGIS estimates with Equation (9) a negative slope β_1 and, hence, a switching point which is economically meaningless. Without taking into account the measurement error, this switching point enters the estimation of Equation (6) with the same weight as any other observation. In contrast, the Bayesian model estimates for these two observations a precision τ_{ik} very close to zero.

The top panel in Figure 2 shows the value of β_1 from Equation (9). The panel in the middle shows the estimated precision τ_{ik} from Equation (5). Comparing both panels, we see that the B-JOINT estimates are more differentiated. The LOGIS estimates for β_1 are either close to positive or negative infinity, or close to zero. By contrast the B-JOINT estimates for precision τ_{ik} show a more detailed picture of deviation from utility maximising behaviour. For the consistent choices the estimated parameter for τ is rather large with a median value of 57.4. For the inconsistent choices τ covers a range from 1.68 to 47.6.

Selection bias versus measurement error: While the results of B-JOINT are based on the entire dataset, including the inconsistent decision makers, we also estimate B-JOINT-DROP, based on the same model but using only data from consistent decision makers. The right part of Table 3 shows results only for the 59 consistent observations. The comparison of the two models, B-JOINT and B-JOINT-DROP, allows us to decide whether our results are mainly driven by accounting for the measurement error or by avoiding a selection bias. Both models take into account the measurement error. Both models come to substantial effect sizes for $\eta \times p$: -5.16 for B-JOINT-DROP, and -3.64 for B-JOINT. Only B-JOINT avoids the selection bias. Above we have seen that accounting for measurement errors changes the coefficient $\beta_{\eta \times p}$ by 210%. The selection bias, i.e. the difference between B-JOINT and B-JOINT-DROP, affects the effect size by only 30%.

4. Simulation

Should one take into account measurement errors? The above result seems to suggest that such a correction is desirable, but how general is this finding? Here we simulate 100 times a

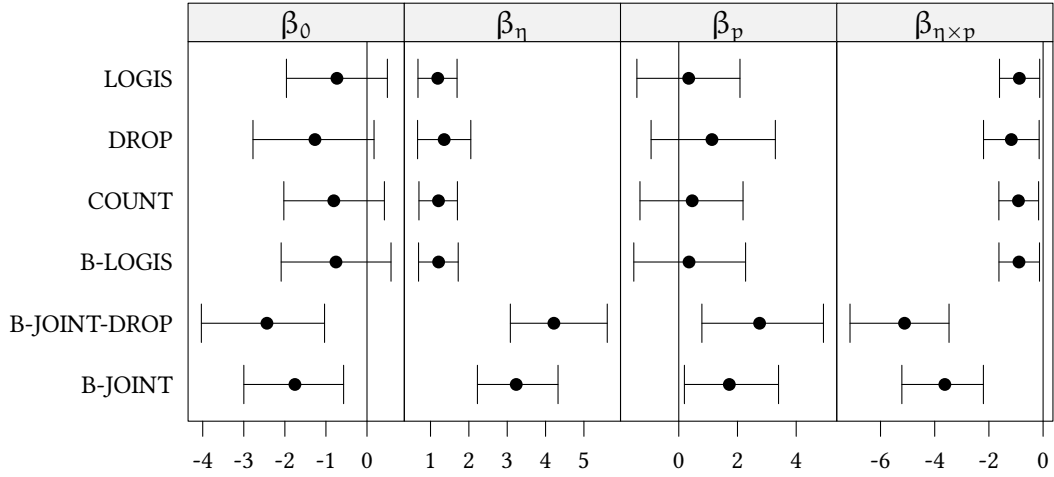


Figure 3: Highest posterior density and confidence intervals for Equation (6).

The figure shows 95% confidence intervals for the mixed effects model based on LOGIS, DROP and COUNT estimates for p_{ik}^c . The figure also shows 95% HPD intervals for three specifications of the Bayesian model: B-LOGIS, which is a replication of the B-LOGIS model in the Bayesian framework based on consistent choices only, B-JOINT-DROP, the joint model based on only consistent choices, and B-JOINT, the joint model for all choices.

sample that is similar to the one we studied above. Each sample has a size of 100 participants which come in 25 groups.

Behaviour in the risk task and in the public good game follows Equations (5) and (11). The parameters of the regression are random and in the same order of magnitude as in our experiment: $\beta_l \sim \mathcal{N}(0, 2)$ for $l \in \{0, \eta, p, \eta \times p\}$. The random effects have a similar variance: $\nu_k \sim \mathcal{N}(0, \sqrt{1/5})$, $\nu'_{ik} \sim \mathcal{N}(0, \sqrt{2/7})$, $\epsilon_{ikt} \sim \mathcal{N}(0, \sqrt{10})$. The risk aversion also follows a distribution similar to the one in our experiment: $p_{ik}^c \sim \mathcal{B}(6.98, 3.63)$, $\tau_{ik} \sim \Gamma(0.847, 0.2)$.

For each of the 100 simulations we obtain an estimate for the coefficients of Equation (6). Here we are specifically interested in $\beta_{\eta \times p}$. Figure 4 shows for both methods COUNT and B-JOINT quartiles of the difference between the estimates and the true values, $\hat{\beta}_{\eta \times p} - \beta_{\eta \times p}$. We see that B-JOINT performs fairly well. The difference $\hat{\beta}_{\eta \times p} - \beta_{\eta \times p}$ is close to zero. The estimates of COUNT are clearly biased. They are too large in the negative and too small in the positive domain. This bias is what we should expect if measurement errors are neglected.

5. Conclusion

We are clearly not alone in recommending that measurement errors should be taken into account in experimental research. Researchers might have the impression, however, that accounting for measurement errors with repeated measurements is tedious and expensive. Many researchers seem to assume the problem away.

The aim of this paper is to convince the experimental community that it makes sense to

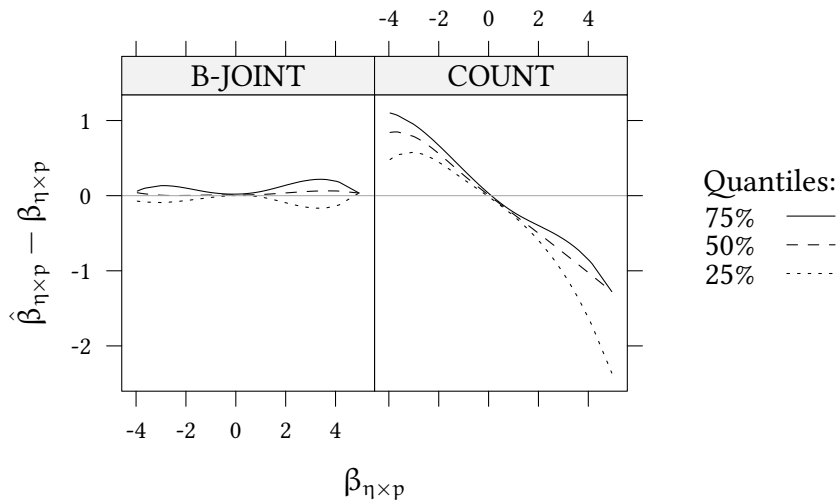


Figure 4: Simulation results: Bias

The figure shows 25%, 50%, and 75% quantiles of a B-spline (df=5) for the bias $\hat{\beta}_{\eta \times p} - \beta_{\eta \times p}$ and for the different models.

take measurement errors seriously, and that it is possible to correct for the resulting bias in a straightforward and simple way, without making the experimental design more complicated.

Here we take advantage of the fact that a single Holt and Laury task is, actually, a multiple price list task, i.e. the task already contains multiple measures. The fact that the Holt and Laury task asks each participant to take multiple risky choices is not a nuisance. It enables the researcher to assess the precision of his or her instrument.

References

- Adcock, R. J. (1877). “Note on the Method of Least Squares”. In: *The Analyst* 4.6, pp. 183–184.
- Arminger, Gerhard and Bengt O. Muthén (1998). “A Bayesian Approach to Nonlinear Latent Variable Models Using the Gibbs Sampler and the Metropolis-Hastings Algorithm”. In: *Psychometrika* 63.3, pp. 271–300.
- Bayarri, M. J. and J. O. Berger (2004). “The interplay of Bayesian and frequentist analysis”. In: *Statistical Science* 19.1, pp. 58–80. DOI: 10.1214/088342304000000116.
- Beauchamp, Jonathan P., David Cesarini and Magnus Johannesson (2017). “The psychometric and empirical properties of measures of risk preferences”. In: *Journal of Risk and Uncertainty* 54.3, pp. 203–237. DOI: 10.1007/s11166-017-9261-3.
- Dellaportas, Petros and David A. Stephens (1995). “Bayesian Analysis of Errors-in-Variables Regression Models”. In: *Biometrics* 51.3, pp. 1085–1095.
- Eckel, Catherine C. and Philip J. Grossman (2008). “Forecasting risk attitudes: An experimental study using actual and forecast gamble choices”. In: *Journal of Economic Behavior & Organization* 68.1, pp. 1–17.

- Engel, Christoph (2014). "Social preferences can make imperfect sanctions work: Evidence from a public good experiment". In: *Journal of Economic Behavior & Organization* 108.C, pp. 343–353.
- Fehr, Ernst and Simon Gächter (2000). "Cooperation and Punishment in Public Goods Experiments". In: *American Economic Review, American Economic Association* 90.4, pp. 980–994.
- Fischbacher, Urs (2007). "z-Tree: Zurich Toolbox for Ready-made Economic Experiments". In: *Experimental Economics* 10.2, pp. 171–178.
- Florens, J. P., M. Mouchart and J. F. Richard (1974). "Bayesian inference in Error-in-Variables Models". In: *Journal of Multivariate Analysis* 4, pp. 419–452.
- Friedman, Daniel, R. Mark Isaac, Duncan James and Shyam Sunder (2014). *Risky Curves: On the Empirical Failure of Expected Utility*. Santa Cruz Department of Economics, Working Paper Series qt87v8k86z. Department of Economics, UC Santa Cruz.
- Gelman, Andrew and Donald B. Rubin (1992). "Inference from Iterative Simulation Using Multiple Sequences". In: *Statistical Science* 7.4, pp. 457–472. DOI: 10.1214/ss/1177011136.
- Gillard, Jonathan (2010). "An overview of linear structural models in errors in variables regression". In: *REVSTAT–Statistical Journal* 8.1, pp. 57–80.
- Gillen, Ben, Erik Snowberg and Leeat Yariv (2015). "Experimenting with Measurement Error: Techniques with Applications from the Caltech Cohort Study". NBER Working Paper #21517.
- Greiner, Ben (2004). "An Online Recruitment System for Economic Experiments". In: *Forschung und wissenschaftliches Rechnen*. Ed. by Kurt Kremer and Volker Macho. Vol. 63. GWDG Bericht. Ges. für Wiss. Datenverarbeitung. Göttingen, pp. 79–93.
- Halevy, Yoram (2007). "Ellsberg Revisited: An Experimental Study". In: *Econometrica* 75.2, pp. 503–536.
- Harrison, Glenn W., Eric Johnson, Melayne M. McInnes and E. Elisabet Rutström (2005). "Temporal stability of estimates of risk aversion". In: *Applied Financial Economics Letters* 1.1, pp. 31–35.
- Hey, John D. and Chris Orme (1994). "Investigating Generalizations of Expected Utility Theory Using Experimental Data". In: *Econometrica* 62.6, pp. 1291–1326.
- Holt, Charles A. and Susan K. Laury (2002). "Risk Aversion and Incentive Effects". In: *The American Economic Review* 92.5, pp. 1644–1655.
- Kass, Robert E (2011). "Statistical inference: The big picture". In: *Statistical science: a review journal of the Institute of Mathematical Statistics* 26.1, p. 1.
- Kimball, Miles S, Claudia R Sahm and Matthew D Shapiro (2008). "Imputing Risk Tolerance From Survey Responses". In: *Journal of the American Statistical Association* 103.483, pp. 1028–1038. DOI: 10.1198/016214508000000139.
- Loomes, Graham (2005). "Modelling the Stochastic Component of Behaviour in Experiments: Some Issues for the Interpretation of Data". In: *Experimental Economics* 8.4, pp. 301–323. DOI: 10.1007/s10683-005-5372-9.
- Neyman, Jerzy and Elizabeth L. Scott (1948). "Consistent Estimates Based on Partially Consistent Observations". In: *Econometrica* 16.1, pp. 1–32.
- Niederle, Muriel and Lise Vesterlund (2007). "Do Women Shy Away from Competition? Do Men Compete Too Much?" In: *The Quarterly Journal of Economics* 122.3, pp. 1067–1101.

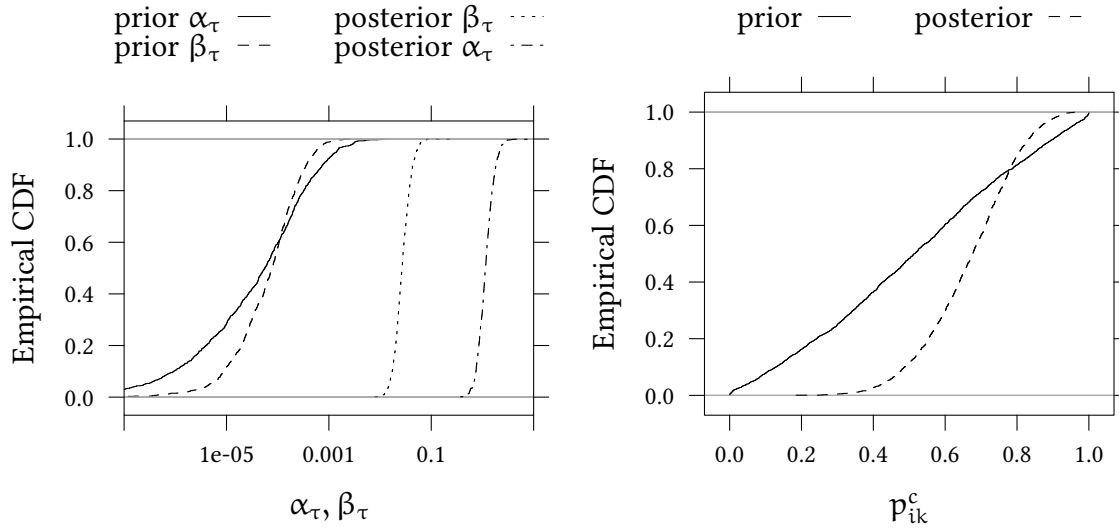


Figure 5: Posteriors for α_τ , β_τ and p_{ik}^c

- Polasek, Wolfgang and Andreas Krause (1993). “Bayesian regression model with simple errors in variables structure”. In: *The Statistician* 42, pp. 571–580.
- R Development Core Team (2018). *R: A Language and Environment for Statistical Computing*. ISBN 3-900051-07-0. R Foundation for Statistical Computing, Vienna, Austria.
- Rabe-Hesketh, Sophia, Anders Skrondal and Andrew Pickles (2004). “Generalized multilevel structural equation modeling”. In: *Psychometrika* 69.2, pp. 167–190. DOI: 10.1007/BF02295939.
- Ross, L., R.E. Nisbett and M. Gladwell (2011). *The Person and the Situation: Perspectives of Social Psychology*. Pinter & Martin Limited.
- Sahm, Claudia R. (2012). “How Much Does Risk Tolerance Change?” In: *Quarterly Journal of Finance (QJF)* 2.04, pp. 1–38. DOI: 10.1142/S201013921250019X.
- Solari, Mary E. (1969). “The ”Maximum Likelihood Solution” of the Problem of Estimating a Linear Functional Relationship”. In: *Journal of the Royal Statistical Society. Series B (Methodological)* 31.2, pp. 372–375.

A. Posteriors for α_τ , β_τ and p_{ik}^c

Figure 5 shows the prior and posterior distribution for p^c and for the parameters $\alpha_\tau = m^2/d^2$, $\beta_\tau = m/d^2$ which determine the distribution of τ_{ik} . In Equation (15) we assume p_{ik}^c follows a Beta distribution with parameters α_c and β_c following a Gamma distribution (so that a priori p_{ik}^c follows an almost uniform distribution). The median of the posterior parameters are $\alpha = 7.89$ and $\beta = 3.95$, i.e., as we also see in the Figure, participants do avoid the extreme values of p^c and, not surprisingly, are more risk averse than risk loving.

In Equation (16) we assume that the precision τ_{ik} is drawn from a Gamma distribution. The median of the posterior shape parameter of this distribution is $\alpha = 1.16$ and the median of the posterior rate parameter is $\beta = 0.027$. Figure 6 shows the posterior distribution of τ_{ik} as well

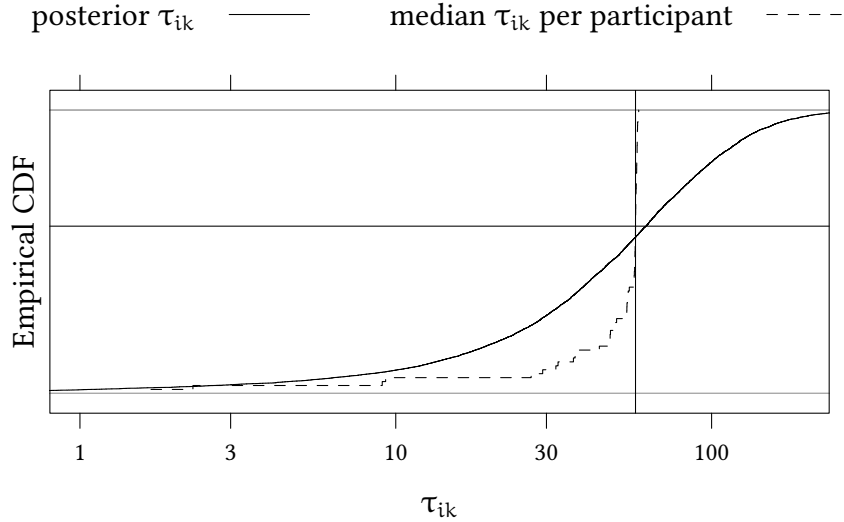


Figure 6: Precision of choices for τ_{ik}

The solid line show the posterior distribution of τ_{ik} as in Equation (16). The dotted line shows the distribution of the median of τ_{ik} taken for each participant.

as the median values of τ_{ik} for the individual participants. Conceptually, this is not entirely trivial. Often we assume that “consistent” choices are infinitely precise, i.e. $\tau = \infty$. However, if some choices, here 18% of all participants, are inconsistent, i.e. contain a substantial lack of precision ($1.68 \leq \tau \leq 47.6$), it would be foolish to assume that the remaining 82% choices are infinitely precise.

How can we assess the precision of choices? In Figure 6 we see how the estimator uses the 18% inconsistent observations as a handle to estimate the left part of the distribution of τ . On the right side of the distribution the value of 57.4 for the median consistent decision maker results from the discrete steps in the Holt and Laury (2002) task which implies a finite precision for the consistent choices.

B. Instructions

General Instructions In the following experiment, you can earn a substantial amount of money, depending on your decisions. It is therefore very important that you read these instructions carefully.

During the experiment, any communication whatsoever is forbidden. If you have any questions, please ask us. Disobeying this rule will lead to exclusion from the experiment and from all payments.

You will in any case receive 4 € for taking part in this experiment. In the first two parts of the experiment, we do not speak not of €, but instead of Taler. Your entire income from these two parts of the experiment is hence initially calculated in Taler. The total number of Taler you earn during the experiment is converted into € at the end and paid to you in cash, at the rate of

$$1 \text{ Taler} = 4 \text{ Eurocent.}$$

The experiment consists of four parts. We will start by explaining the first part. You will receive separate instructions for the other parts.

Part One of the Experiment In the first part of the experiment, there are two roles: A and B. Four participants who have the role A form a group. One participant who has the role B is allocated to each group. The computer will randomly assign your role to you at the beginning of the experiment.

On the following pages, we will describe to you the exact procedure of this part of the experiment.

Information on the Exact Procedure of the Experiment This part of the experiment has two steps. In the first step, role A participants make a decision on contributions to a project. In the second step, the role B participant can reduce the role A participants' income. At the start, each role A participant receives 20 Taler, which we refer to in the following as the endowment. Role B participants receive 20 points at the start of step 2. We explain below how role B participants may use these points.

Step 1: In Step 1, only the four role A participants in a group make a decision. Each role A member's decision influences the income of all other role A players in the group. The income of player B is not affected by this decision. As a role A participant, you have to decide how many of the 20 Taler you wish to invest in a project and how many you wish to keep for yourself.

If you are a role A player, your income consists of two parts:

1. the Taler you have kept for yourself ("income retained from endowment")
2. the "income from the project". The income from the project is calculated as follows:

Your income from the project = 0.4 times the total sum of contributions to the project

Your income is therefore calculated as follows:

$$(20 \text{ Taler} - \text{your contribution to the project}) + 0.4 * (\text{total sum of contributions to the project}).$$

The income from the project of all role A group members is calculated according to the same formula, i.e., each role A group member receives the same income from the project. If, for example, the sum of the contributions from all role A group members is 60 Taler, then you and all other role A group members receive an income from the project of $0.4 * 60 = 24$ Taler. If the role A group members have contributed a total of 9 Taler to the project, then you and all other role A group members receive an income from the project of $0.4 * 9 = 3.6$ Taler.

For every Taler that you keep for yourself, you earn an income of 1 Taler. If instead you contribute a Taler from your endowment to your group's project, the sum of the contributions to the project increases by 1 Taler and your income from the project increases by $0.4 * 1 = 0.4$ Taler. However, this also means that the income of all other role A group members increases by 0.4 Taler, so that the total group income increases by $0.4 * 4 = 1.6$ Taler. In other words, the other role A group members also profit from your own contributions to the project. In turn, you also benefit from the other group members' contributions to the project. For every Taler that another group member contributes to the project, you earn $0.4 * 1 = 0.4$ Taler.

Please note that the role B participant cannot contribute to the project and does not earn any income from the project.

Step 2: In Step 2, only the role B participant makes decisions. As role B participant, you may reduce or maintain the income of every participant in Step 2 by distributing points.

At the beginning of Step 2, the four role A participants and the role B participant are told how much each of the role A participants has contributed to the project.

As a role B player, you now have to decide, for each of the four role A participants, whether you wish to distribute points to them and, if so, how many points you wish to distribute to them. You are obliged to enter a figure. If you do not wish to change the income of a particular role A participant, please enter 0. Should you choose a number greater than zero, you reduce the income of that particular participant. For each point that you allocate to a participant, the income of this participant is reduced by 3 Taler.

The total Taler income of a role A participant from both steps is hence calculated using the following formula:

$$\text{Income from Step 1} - 3 * (\text{sum of points received})$$

Please note that Taler income at the end of Step 2 can also be negative for role A participants. This can be the case if the income-subtraction from points received is larger than the income from Step 1. However, the role B participant can distribute a maximum of 20 points to all four role A members of the group. 20 points are the maximum limit. As a role B participant, you can also distribute fewer points. It is also possible not to distribute any points at all.

If you have role B, please state your reasons for your decision to distribute (or not to distribute) points, and why you distributed a particular number of points, if applicable. In doing this, please try to be factual. Please enter your statement in the corresponding space on your screen. You have 500 characters max. to do this. Please note that, in order to send your statement, you will have to press “Enter” once each time. As soon as you have done this, you will no longer be able to change what you have written.

The income of the role B participant does not depend on the income of the other role A participants, nor on the income from the project. For taking part in the first part of the experiment, he or she receives a fixed payment of

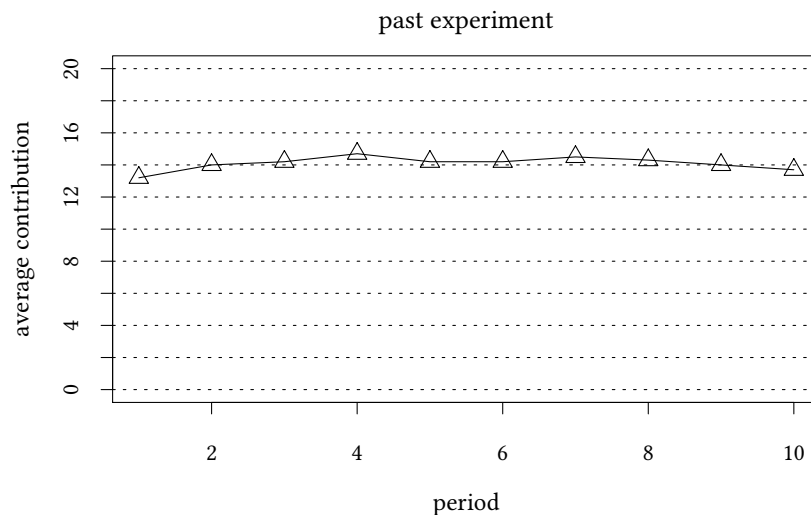
1 €.

In addition, the role B participant receives the sum of 0.01 € for each point that he or she did not distribute. Once all participants have made their decisions, your screen will show your income for the period and your total income so far.

After this, the first part of the experiment ends. You will then be told what your payment is for this part of the experiment. Hence, you will also know how many points you and all other participants have been given by player B.

Experiences from an Earlier Experiment For your information, we give you the following graph, which tells you the average contributions made in a very similar experiment that was conducted in this laboratory.

In this experiment, too, there were groups of 4 role A participants and one role B participant each. The role A participants’ income was calculated in exactly the same way. The experiment had 10 equal periods. The role B participant also had 20 points at his disposal in each period. At the end of each period, the role A participants were told how much each of the other participants had contributed and how the role B participant had reacted to this.



Part Two of the Experiment The second part of the experiment consists of 10 repetitions of the first part. Throughout the entire second part, all participants keep the role they had in the first part of the experiment.

The computer randomly rematches the groups of four in every period. In each period, the computer randomly assigns a role B participant to each group.

As a reminder: In each period, each role A participant receives 20 Taler, which may be contributed to the project entirely, in part, or not at all. For each period, calculating the income from the project for the role A participants in a group happens in exactly the same way as it did in the first part of the experiment. In each period, each role B participant receives 20 points, which may be used to reduce the income of the players A in the group. For each point that a role A participant receives in a period, 3 Taler are subtracted. For each point that a role B participant does not use, he or she is given the sum of 0.01 €. In addition to the income from the points retained, each role B participant receives a flat fee of 10 € for participating in this second part of the experiment.

At the beginning of Step 2 of each period, the four role A participants and the role B participant are told how much each of the role A participants contributed to the project.

Please note that the groups are rematched anew in each period.

After each period, you are told about your individual payoff. You are therefore also informed how many points you and the other participants have been assigned by the role B participant.

Part Three of the Experiment We will now ask you to make some decisions. In order to do this, you will be randomly paired with another participant. In several distribution decisions, you will be able to allocate points to this other participant and to yourself by repeatedly choosing between two distributions, 'A' and 'B'. The points you allocate to yourself will be paid out to you at the end of the experiment at a rate of 500 points = 1 €. At the same time, you are also randomly assigned to another participant in the experiment, who is, in turn, also able to allocate points to you by choosing between distributions. This participant is not the same participant as the one to whom you have been allocating points. The points allocated to you are also credited to your account. The sum of all points you have allocated to yourself and those allocated to you by the other participant are paid out to you at the end of the experiment at a rate of 500 points = 1 €. Please note that the participants assigned to you in this part of the experiment are not the members of your group from the preceding part of the experiment. You will therefore be dealing with other participants.

The individual decision tasks will look like this:

| Possibility A | | Possibility B | |
|---------------|---|---------------|---|
| Your points: | The points of the participant of the experiment allocated to you: | Your points: | The points of the participant of the experiment allocated to you: |
| 0 | 500 | 304 | 397 |

In this example: If you click 'A', you give yourself 0 points and 500 points to the participant allocated to you. If you click 'B', you give yourself 304 points and 397 points to the participant allocated to you.