A Note on an Alternative Approach to Experimental Design of Lottery Prospects

Balcombe, Kelvin and Fraser, Iain

University of Reading, University of Kent

January 2024

A Note on an Alternative Approach to Experimental Design of Lottery Prospects

Kelvin Balcombe
University of Reading

Iain Fraser*
University of Kent

8th January 2024

***Corresponding Author:**
Professor Iain Fraser
School of Economics
University of Kent
Canterbury
Kent, CT2 7NP
Tel: 01227 823513
Email: i.m.fraser@kent.ac.uk

Acknowledgements: This research was part funded by British Academy/Leverhulme Small Research Grants (SRG\170874) "Improving the estimates of attitudes towards risk."

Conflict of Interest
The authors declare that they have no conflict of interest.

A Note on an Alternative Approach to Experimental Design of Lottery Prospects

Abstract

We introduce an alternative approach to lottery prospects experimental design aimed at collecting experimental data for parametric estimation of the cumulative form of Prospect Theory (PT). Our approach incorporates two fundamental principles: ensuring that all tasks provide valuable information and avoiding redundancy among tasks. These principles mean that we avoid the construction of lottery prospects that duplicate information within the set of tasks generated. The methodological approach that we have designed ensures that each lottery pair is non-redundant in an informational sense. This means that the set of lottery tasks generated can help to improve the effectiveness of data collection when estimation of preference parameters is the main research objective. In this note, we describe our approach to experimental design in detail.

Key Words: Experimental Design; Lotteries; Risk and Uncertainty; Prospect Theory.

JEL: C11; C52; D81.

1 Introduction

In this note, we introduce an alternative approach to lottery design that can be used in an experimental setting to collect data that in turn can be used to estimate preference parameters for cumulative Prospect Theory (PT) (Tversky and Kahneman, 1992). Our aim is to develop an alternative approach to experimental design that can enable the effective estimation of preference parameters. There exists a large literature estimating preference parameters using experimental data (e.g., Stott, 2006; Booij et al., 2010; Nilsson et al., 2011; Balcombe and Fraser, 2015; Balcombe et al., 2019; Kpegli, et al., 2023) Estimation of preference parameters is one reason As such this research contributes to an important issue, that is the formulation and resulting effectiveness and efficiency of a lottery design that underpins any experimental data collection (Loomes and Sugden, 1998; Hey 2001; Cavagnaro et al., 2013a,b; Moffat, 2015).

It is well understood that any experimental design will have an impact on experimental data collection and potentially influence the preference parameters parametrically recovered from the data. If you “assume” (your prior) any key PT parameter is likely to be similar in value to those originally reported by Tversky and Kahneman (1992) and you incorporate this into your experimental design (e.g. when simulating lotteries) it is then highly likely that your experiment will yield PT parameters estimates within the "expected range". Indeed, if you develop experimental designs employing simulation methods that have "tight" bounds on your prior beliefs, your resulting set of lotteries will likely reflect your priors. Consequently, how prior beliefs about parameters shape the experimental design needs to be explicitly recognized and described by researchers when estimating preference parameters with such data.

Here we address the basic question of how should researchers select the actual set of lotteries to be used in their experiment. We address this question by introducing a simple but logical approach to experimental design. Our approach is based upon two basic principles. First, each task a respondent undertakes needs to be "informative". This simply means that respondents who have different preferences will select different options in a lottery choice task. Second, no task should be rendered redundant by any other task. The most obvious example is that a task should not be repeated. More generally, it means that one task should not be able to be used to predict responses to other tasks. By employing these two principles within an algorithm, we can construct lotteries that are informative and have low pairwise redundancy defined by using an entropy measure.

2 The Design of Lottery Tasks

2.1 Common Design Issues

When a researcher sets out to design an experiment such as a binary lottery or multiple price list (MPL), that can be used to estimate PT parameters there are many issues to consider including: what statistical method/criteria to employ to derive lotteries; the number of lotteries to generate (ie, the super-set); selecting the sub-set to be used from the super-set; evaluation of the sub-set selected; inclusion of a sure thing option; two or more lottery outcomes; text, graphics or a mix presentation of the task; and purely hypothetical lotteries, or do you offer real incentives. Several of these issues have received far more attention than others. For example, there is a large literature on how risk preferences can be elicited and the mechanism to be used (Charness et al., 2013; Crosetto and Filippin, 2016) with the choice of method being highly dependent on the researcher's goals. Many papers have employed either lists, ladders, or a series of discrete lottery choices following the procedures introduced by Eckel and Grossman (2002, 2008) and the MPL approach introduced by Holt and Laurey (2002). There is a large research literature comparing elicitation methods (e.g., Charness et al., 2013; Crosetto and Filippin, 2016; Drichoutis and Lusk, 2016; Pedroni et al., 2017; Freeman and Mayraz, 2019; Holzmeister and Stefan, 2021).

Another topic that has attracted a lot of attention is the number of tasks undertaken during the experiment. It is assumed that offering respondents ever more tasks enables the recovery of better parameter estimates, such that the number of tasks employed can be large. For example, Rieskamp (2008) gave 30 participants 180 pairs of gambles with 60 in each domain (Gain, loss, and mixed). Others (e.g. Hey and Orme, 1994; Hey, 2001; Stott, 2006) have opted for 100 tasks. More recently Frydman and Jin (2022) have asked respondents to complete 600 tasks. However, as noted by Hey (2001): "*The more questions the better as long as tiredness does not set in.*" (p.7). Furthermore, simply increasing the number of tasks given to respondents need not necessarily improve the accuracy of model estimates. This is because, without careful consideration, a researcher might simply use informationally redundant lotteries. Moreover, increasing the number of tasks may induce respondent fatigue or reduce engagement and degrade the quality of the data collected impacting the preference parameter estimates recovered.

There are also examples of researchers recycling experimental tasks which presumes that the existing experimental design is both sound and appropriate for the task in hand. For example, the lottery design developed by Binswanger (1980) has been widely used (e.g. Bauer et al., 2012; Chowdhury et al., 2022). Another is that developed and employed by Tanaka et al. (2010) that has been re-used by many researchers (e.g. Liu and Huang, 2013; Ward and Singh, 2015; Bougherara et al., 2017). You also find examples of researchers combining existing sets of experimental designs (e.g. Andersen et al., 2018; Murphy and ten Brincke, 2018).

2.2 Efficiency of Experimental Designs

Within the literature, there are often extensive descriptions of how experiments have been implemented and how they deal with the issues already noted. However, there is frequently far less detail given about how the lotteries or choice tasks were generated and selected. For example, Andersen et al. (2018) indicate that they carefully selected the lottery tasks employed such that they could econometrically identify the structural model of interest. They also detail and explain very carefully the sources of all of the lotteries employed, but they do not give any more information as to how the selection of the lotteries may be efficient. Indeed the tasks employed by Andersen et al. (2018) draw on some prominent experimental designs that appeal to theory to justify the tasks generated. For example, Hey (2001) motivates the choice of lotteries using Marschak–Machina Triangles (MMT) as do Harrison and Swarthout (2016), following the approach taken by Loomes and Sugden (1998). While this approach provides an excellent platform for generating tasks, it can be implemented in many ways, and the performance of the final design in terms of estimating preference parameters is not clear. Both Cavagnaro et al. (2013a) and Harrison and Ng (2016) provide extended and insightful discussions regarding MMTs and experimental design.

In general, the statistical efficiency of lottery designs has attracted relatively little attention compared to other experimental areas in economics, such as stated preference research (e.g., Johnson et al., 2017). Why is this the case? Maybe, in some part, this is because there was a historical tendency to think about risk preferences in a non-stochastic/deterministic way. However, it also stems from the wide range of candidate models for risk preferences which seriously complicates attempts to generate lottery designs using statistical criteria.

Importantly, there is no such thing as an "efficient design" per se. A design that is highly efficient for estimating the parameters of one model may perform very poorly for estimating the parameters of another (possibly more correct) model. For example, if a researcher assumes that key PT parameters are going to be in a range similar to those originally reported by Tversky and Kahneman (1992), a set of lotteries can be generated that is capable of estimating parameters in that neighborhood, yet might be highly uninformative about parameter values that lie far outside this range. Somewhat uncomfortably, this also implies the possibility that the apparent superior performance of one model over another may not be due to its inherent superiority but due to the particular set of tasks within the experimental design. This issue has also been discussed in relation to MMTs and the location of lottery pairs (Harrison and Ng, 2016).

To date, only a few papers have considered the statistical properties of experimental designs in regard to eliciting risk preference parameters (e.g. Müller and Ponce de Leon, 1996; Moffatt, 2007, 2015; Cavagnaro et al., 2013a). Müller and Ponce de Leon (1996) specifically address the issue of experimental design concerning risk elicitation, although they appear to focus on the functional form and related issues. Moffatt (2007) discusses how to employ D-optimal designs (an approach commonly employed in the discrete choice experiment literature) but does not implement an approach due to the complexity of implementation involved. Moffatt (2015) also discusses the possibility of employing D-optimal designs but explains why they are computationally too demanding.

A very different philosophical approach to experimental design is used by Cavagnaro et al. (2013a). They employed a design approach that is referred to as adaptive design optimization (ADO). ADO uses an algorithm to dynamically select the choice tasks, not in advance of the experiment but during the experiment. Employing Bayesian inference, ADO requires that researchers explicitly specify priors for models and parameters such that after every choice is made the subsequent choice offered to an individual respondent is informative.[1] As a result, ADO can yield efficient designs. However, as noted by Sloman et al. (2023) the choice of the priors has a significant effect on the performance of ADO. They also report that the selection of priors has significantly less impact on parameter estimates but that model selection can be significantly affected. There is much to like about the ADO approach, but it does require the researcher to have strong priors in advance of data collection.

Within the wider decision theory literature there are also papers that examine the issue of efficient experimental design. An example is the paper by Broomell and Bhatia (2014) that has recently been used by Olschewski et al. (2021). Broomell and Bhatia (2014) develop a criterion to rank decision sets that allows for parameter discrimination. This approach provides important insights into the ease with which different PT parameters can be identified from experimental designs. This is an important observation as insufficient attention is paid to why a design method was chosen and how well it would identify the parameters of a given model. When rationalizations are given, these are often in terms of identifying tasks that might discriminate between models or elicit choices that contradict a given model. We contend that there is an under-appreciation of the necessary amount of information required to identify the

[1] The implementation of ADO has been enabled by the development of a python package ADOpy by Yang et al. (2021).

parameters of interest for flexible models derived from PT, particularly if there is respondent heterogeneity.

2.3 Task Complexity and the Sure-Thing Option

An important issue to be aware of when undertaking any experimental design is the inherent degree of task complexity. When undertaking experimental design from a purely statistical perspective the degree of task complexity confronting respondents also needs to be appreciated. Researchers need to strike a balance between making tasks too easy or giving more complex tasks that, in theory, might be more informative in regard to revealing information but respondents find challenging and therefore fail to engage properly with the experiment.

The role of complexity in survey designs has been examined at length in various research areas, including stated preference discrete choice experiments (e.g. Pfeffier et al., 2014; Johnson et al., 2017; Regier et al., 2014) and lottery/prospect designs. For example, complexity has been considered in the lottery design literature in terms of similarity of prospect pairs (e.g. Buschena and Zilberman, 1999; Buschena and Atwood, 2011). The issue of complexity has also been examined with regard to preference reversals (Loomes and Pogrebna (2017)[2] and respondent performance by Charness et al. (2018) who note that complexity and the structure of a risk elicitation mechanism impact measured risk preferences. Amador-Hidalgo et al.(2021) also note that commonly employed lottery tasks, such as MPLs frequently result in a significant proportion of inconsistent choices. In terms of explaining inconsistent choices Amador-Hidalogo et al. (2021) report findings in keeping with Andersson et al. (2016). As task complexity increases, that is the probability calculations are more complicated, then the number of inconsistent choices increases. Inconsistent choices also increase as the realization of the probabilities of the two payoffs get closer because the computation required to identify the preferred option becomes harder. Amador-Hidalgo et al. (2021) suggest one way to deal with this problem would be to employ more tasks but this then runs up against the fatigue issue already noted. Andersson et al. (2016) conclude that in an effort to reduce problems with inconsistent tasks that the following is required when designing choice tasks: "*...the need to use balanced experimental elicitation designs (e.g., several MPLs with varying switch points for given risk preferences.*" (p. 1131). Although there is obvious merit in this suggestion, it is somewhat vague and not entirely clear what it means when it comes to practical implementation. They also suggest what they refer to as a balanced design, but this is not a useful criteria or practical solution to experimental task design.[3]

One way in which researchers have sought to devise tasks that help reduce

[2] The issue of preference reversals is a long-standing research theme eg., Lichtenstein and Slovic (1971) and Grether and Plott (1979). However, examination of "reversal rates" generally requires the repetition of tasks, an approach not recommended if the goal is to maximise the information elicited about preferences within a fixed number of tasks.

[3] The impact of task complexity has also been discussed in relation to the legitimacy of rank dependency, see Bernheim and Sprenger, (2020) and Bernheim et al. (2022).

complexity is by limiting the number of tasks in sets of lotteries to those that only require prospects with two or three payoffs or where one of the options is a sure-thing (e.g. Bruhin et al., 2010; Falk et al., 2018; l'Haridon and Vieider, 2019). By employing a sure-thing in an experimental design this can give rise to what Tversky and Kahneman (1981) coined as certainty bias. This is a phenomenon where respondents chose options that have payoffs that are certain but cannot be explained by expected utility theory (EUT). Kahneman and Tversky (1986) view the certainty effect as a framing bias, but it can potentially be understood as a phenomenon that arises because respondents choose options that they can more easily understand (albeit because of framing). Interest in certainty bias is ongoing with Zilker and Pachur (2021) explaining how probability weighting within PT can vary the valuation placed upon the risky option but not the safe/sure option. They also contend that it may well be how respondents allocate their attention to a task that gives rise to the observed certainty effect. Frydman and Jin (2022) also examine decision making when confronted with one risky choice and a sure thing. They consider how changing the range of payoffs effects decision making, such that with a higher range of values, a treatment they refer to as the high-volatility setting, the likelihood of selecting the sure thing decreases. What they take this to mean is that when a respondent's perceptions are noisier and this occurs in high volatility settings this in turn generates noisier (inconsistent) choices. Frydman and Jin (2022) express this finding as follows:

"..., we provide evidence consistent with the hypothesis that diminishing sensitivity to payoffs arises in part from an optimal allocation of perceptual resources." (p. 166).

2.4 Summary

This brief review of the literature indicates that the existing experimental literature that has set out to examine risk preferences, frequently with a focus on PT has generally paid minimal attention to issues relating to experimental design efficiency as it relates to parametric derivation of preference parameters. Few studies discuss the relative efficiency of an experimental design relative to a defined criterion. This issue is explicitly considered by Moffat (2007, 2015), but only Cavagnaro et al. (2013 a,b) and to a lesser extent Broomell and Bhatia (2014) have developed methods to evaluate the efficiency of a set of lotteries or prospects used to statistically derive key PT parameters. Although both methods have merit there remains scope to re-examine this issue from a different philosophical perspective as we do in this paper. There is also an extensive discussion on task complexity and design choice, such as the use of the sure-thing, but this has not been explicitly linked to how better to undertake the design of experimental tasks. This is surprising given the extensive literature that exists in other areas of experimental design such as stated preference research. It is these two issues that we address in the approach to design we present in this note.

3 Experimental Design

3.1 In Principle

Our lottery tasks are presented to respondents as a series of discrete choices. To reduce task complexity, we have one of the options as a sure-thing, plus either a two-payoff or three-payoff lottery. We include three payoff options, as well as two payoff probabilities, because it gives greater information about the nature of the probability weightings.

The development of the statistical procedure to generate the set of lotteries employed to estimate PT models is based on two principles:

Principle 1: Each task should be "informative". A task is "non-informative" if the same option would be chosen regardless of preferences. That is, if two individuals with very different preferences are likely to make the same choice, then that task is not informative about the preferences of those individuals. By contrast, an informative task is likely to reveal different choices by individuals with different preferences.

Principle 2: Any task should not be rendered redundant by any other task (pairwise redundancy). The most obvious example is that tasks should not be repeated. However, more generally, the answer to one task should not be able to predict the answer to any other task across the range of possible preferences. This principle relates directly to a limitation of the approach introduced by Broomell and Bhatia (2014). Their approach relies on the assumption that each choice task can be treated as independent of all others. However, this requirement is important as it ignores the dependence in choice tasks in terms of the amount of information revealed.

Using these principles, we can construct a set of lotteries that are highly informative and have low pairwise redundancy. A deeper approach would not only look at the pairwise redundancy of tasks but seek to ensure that each task was informative relative to the entire set of tasks. However, this is a difficult computational problem, therefore we do not attempt to operationalize this in our design. We also note that our approach to experimental design is philosophically different from previous approaches in that we focus on the informational content of the lotteries in an effort to generate a set that will enable effective recovery of model and key parameter estimates.

These principles can be formalized using Bayesian inference that seeks to estimate a ‘posterior’ distribution for the parameters in question. The posterior is the distribution given the choices of respondents and is constructed from the data along with a prior distribution. The more informative this posterior distribution is, the better we can make inferences about the parameters of interest. A set of tasks can be chosen that will deliver a posterior distribution with low entropy, or equivalently, high Kullback-Leibler divergence from a uniform distribution. A flat distribution does not necessarily translate into a non-informative distribution. However, for the parameters used here there should be a correspondence between the flatness and non-informative nature of the distribution, since none of them involve variance parameters of the likelihood, where noninformativeness would normally be modelled using a diffuse gamma.

Under diffuse priors, maximum likelihood estimates are similar to Bayesian ones. Thus, it should be equally applicable to sets designed for Classical estimation procedures. A formal description of our statistical procedure is provided below. The essential components are as follows:

- define a set of T prospect pairs as $\mathbb{P} = (\mathcal{P}_1, \mathcal{P}_2, ..., \mathcal{P}_T)$;
- define the associated outcomes (choices) as $y = (y_1, y_2, y_3, ..., y_T)$; and
- define the preference parameters as ϕ.

The preference parameters are those needed for the standard power form of the PT model $\phi = (\alpha, \lambda, \beta, \gamma, \delta, \rho)$, where α, λ,and β are the parameters that determine the conventional power value functions, γ and δ are the parameters that govern the probability warping of the probability weighting functions and ρ is the parameter that determines the noise in the model.

The entropy of the posterior $f(\phi|y, \mathbb{P})$ based on the logit model employing a PT structure for its systematic utility and a prior for the preferences ϕ is:

$$\xi(\Phi|Y, \mathbb{P}) = -\int_{\phi \in \Phi} \sum_{y \in Y} f(\phi, y|\mathbb{P}) \ln f(\phi|y, \mathbb{P})\, d\phi \tag{1}$$

Ideally this quantity should be minimized and this quantity can be simulated for a given set of tasks $\mathbb{P}$. However, it is quite intensive for large T and therefore an algorithm to choose $\mathbb{P}$ so as to minimize $\xi(\Phi|Y, \mathbb{P})$ is infeasible. However, it is feasibly simulated for all given tasks $\xi(\Phi|Y_i, \mathcal{P}_i)$ and pairs of tasks $\xi(\Phi|Y_i, Y_j, \mathcal{P}_i, \mathcal{P}_j)$. Thus, if $\xi(\Phi|Y_i, \mathcal{P}_i)$ is large then task i is singularly informative and if $\xi(\Phi|Y_i, Y_j, \mathcal{P}_i, \mathcal{P}_j) - \xi(\Phi|Y_i, \mathcal{P}_i)$ is large, then task j is pairwise informative relative to task i.

Given these statistical measures, we next summarize our design algorithm:

1. Decide on the number of tasks to be completed by respondents (we selected 100);
2. Generate a large number of potential tasks randomly (initially 5,000);
3. Simulate the measures of how singularly informative and pairwise informative the tasks are based on a prior distribution for the parameters of interest (i.e. in the underlying PT model as a logit model);
4. Eliminate tasks that are pairwise uninformative relative to tasks that are more singularly informative (if they fall in a given threshold); and
5. Change the threshold and repeat (**4**) until only 100 tasks remain. The selection of the threshold value can be automated until the desired number of lotteries is derived. Alternatively, the threshold can be used to derive a range of lotteries that differ in terms of number of options for the risky option or a lottery that does not include a sure-thing.

The tasks are generated to have payoffs between -100 and 100 where there could be gain, loss, or mixed prospects. The probabilities were set to be increments of 0.05. from 0 to 1. The priors that we used to generate our PT parameters were bounded $\alpha \sim Uniform(0.05, 1.5)$, $\beta \sim Uniform(0.15, 2)$, $\lambda \sim d(0.33, 3)$, $\gamma_1, \gamma_2, \delta_1, \delta_2 \sim d\,(0.4, 2.5)$ where $d\,(a, b)$ is a shifted beta distribution over the interval (a, b) with a modal value of 1 which covered the consensus regions while giving higher prior modal weight to the expected utility model. We set the logit standard deviation as $\rho^{-1} = 2.5(\rho = 0.4)$. This value was chosen so that a certainty equivalent difference of five would give an approximate 90% probability to the task with a higher certainty equivalent and approximately 98% chance if the certainty equivalent difference was as high as 10. In addition, given the size of our payoffs ($|x| < 50$) the parameters of the value function are set so that $\omega_1 < 10$ and $0.15 < \omega_2 < 1$ which allows for a S–type shape for lower payoffs but CRRA/liking for the higher payoffs. At $\omega_1 = 0$ the function collapses to a standard power form in the respective domain.

3.2 Theoretical Derivations

3.2.1 Basic Model Structure

Assume that we have the t^{th} set of prospect pairs $\{\mathcal{P}_t = (\mathcal{Z}_{a,t}, \mathcal{Z}_{b,t})\}$ and a set of preference parameters ϕ that decide preferences. Our aim is to generate a set of $t = 1, ..., T$ prospect pairs where a choice between each of the prospect pairs is informative in the sense that we are likely to learn more about individuals preferences, conditional on what we know from all the other choices. For this design, we use a model based on the deterministic power value function and a Prelec II probability weighting function with a stochastic component for each prospect pair $\mathcal{P}_t = (\mathcal{Z}_{a,t}, \mathcal{Z}_{b,t})$. We will denote this as $V^*\left(\mathcal{Z}_{a,t}, \theta, \gamma, \delta\right)$ where $\theta = (\alpha, \lambda, \beta)$ such that:

$$\begin{aligned} U_{a_t} &= \rho\left(V^*\left(\mathcal{Z}_{a,t}, \theta, \gamma, \delta\right)\right) + e_{a,t} \\ U_{b_t} &= \rho\left(V^*\left(\mathcal{Z}_{b,t}, \theta, \gamma, \delta\right)\right) + e_{b,t} \end{aligned} \tag{2}$$

where the errors $e_{a,t}$ and $e_{b,t}$ are independent Gumbel distributed errors. The probability that $\mathcal{Z}_{a,t}$ will be chosen over $\mathcal{Z}_{b,t}$ is:

$$\Gamma\left(\mathcal{P}_t, \theta, \gamma, \delta, \rho\right) = \frac{e^{\rho V^*(\mathcal{Z}_{a,t}, \theta, \gamma, \delta) - \rho V^*(\mathcal{Z}_{b,t}, \theta, \gamma, \delta)}}{1 + e^{\rho V^*(\mathcal{Z}_{a,t}, \theta, \gamma, \delta) - \rho V^*(\mathcal{Z}_{b,t}, \theta, \gamma, \delta)}} \tag{3}$$

The parameters of interest are $\phi = (\theta, \gamma, \delta)$.

Let $y_t = 1$ if $\mathcal{Z}_{a,t}$ is preferred to $\mathcal{Z}_{b,t}$ and 0 otherwise. It then follows that the probability distribution for y_t is:

$$f\left(y_t | \phi, \rho, \mathcal{P}_t\right) = \left(\Gamma\left(\mathcal{P}_t, \phi, \rho\right)\right)^{y_t} \left(1 - \Gamma\left(\mathcal{P}_t, \phi, \rho\right)\right)^{1 - y_t} \tag{4}$$

Define a set of T prospect pairs $\mathbb{P}$ and associated outcomes (choices) y as follows:

$$\begin{aligned} y &= (y_1, y_2, .., y_T) \\ \mathbb{P} &= (\mathcal{P}_1, \mathcal{P}_2, ..., \mathcal{P}_T) \end{aligned} \tag{5}$$

and denote the set of all possible outcomes for y as Y. Given independent choices (conditional on ϕ) the conditional distribution of y is:

$$f(y|\phi, \rho, \mathbb{P}) = \prod_{t=1}^{T} f(y_t|\phi, \rho, \mathcal{P}_t) \tag{6}$$

Assume that there is a prior density on the underlying preference parameters ϕ and the error ρ

$$f(\phi, \rho) = f(\phi) f(\rho) \tag{7}$$

In what follows, we assume a strict prior for ρ of the form $f(\rho) = 1$ where $\rho = \rho^*$ and 0 otherwise and ignore it subsequently. Therefore, using Bayes theorem, the posterior distribution of ϕ is $f(\phi|y, \mathbb{P})$ is:

$$f(\phi|y, \mathbb{P}) = \frac{f(y|\phi, \mathbb{P}) f(\phi)}{f(y, \mathbb{P})} \tag{8}$$

The marginal likelihood of y is the integrating constant of the posterior:

$$f(y|\mathbb{P}) = \int_{\phi \in \Phi} f(y|\phi, \mathbb{P}) f(\phi) \, d\phi \tag{9}$$

3.2.2 Entropy Measures

If $f(\phi)$ is continuous over the domain Φ, the differential marginal and conditional entropies can be defined. In each case they will be conditioned on $\mathbb{P}$ (the which is the set of tasks taken to be of a fixed number).

- Entropy of the Prior $f(\phi)$: $\xi(\Phi) = -\int_{\phi \in \Phi} f(\phi) \ln f(\phi) \, d\phi$
- Entropy of the Marginal Likelihood: $\xi(Y|\mathbb{P}) = -\sum_{y \in Y} f(y|\mathbb{P}) \ln f(y|\mathbb{P})$
- Entropy of the Posterior $f(\phi|y, \mathbb{P}) : \xi(\Phi|Y, \mathbb{P}) = -\int_{\phi \in \Phi} \sum_{y \in Y} f(\phi, y|\mathbb{P}) \ln f(\phi|y, \mathbb{P}) \, d\phi$
- Entropy of the Likelihood $f(y|\phi, \mathbb{P}) : \xi(Y|\phi, \mathbb{P}) = -\int_{\phi \in \Phi} \sum_{y \in Y} f(y, \phi|\mathbb{P}) \ln f(y|\phi, \mathbb{P}) \, d\phi$

According to the definitions these four are related

$$\xi(\Phi|Y, \mathbb{P}) = \xi(Y|\Phi, \mathbb{P}) + \xi(\Phi) - \xi(Y|\mathbb{P}) \tag{10}$$

This is shown in Proof A:

Proof A
Using

$$f(\phi|y, \mathbb{P}) = \frac{f(y|\phi, \mathbb{P})}{f(y|\mathbb{P})} f(\phi) \tag{11}$$

$$
\begin{aligned}
\xi\left(\Phi|Y,\mathbb{P}\right) &= -\sum_{y\in Y} f\left(y|\mathbb{P}\right)\int_{\phi\in\Phi} f\left(\phi|y,\mathbb{P}\right)\ln f\left(\phi|y,\mathbb{P}\right)d\phi \\
&= -\sum_{y\in Y} f\left(y|\mathbb{P}\right)\int_{\phi\in\Phi} \frac{f\left(y|\phi,\mathbb{P}\right)}{f\left(y|\mathbb{P}\right)} f\left(\phi\right)\ln\left[\frac{f\left(y|\phi,\mathbb{P}\right)}{f\left(y|\mathbb{P}\right)} f\left(\phi\right)\right]d\phi \\
&= -\sum_{y\in Y} f\left(y|\mathbb{P}\right)\int_{\phi\in\Phi} \frac{f\left(y|\phi,\mathbb{P}\right)}{f\left(y|\mathbb{P}\right)} f\left(\phi\right)\left[\ln f\left(y|\phi,\mathbb{P}\right)-\ln f\left(y|\mathbb{P}\right)+\ln f\left(\phi\right)\right]d\phi \\
&= -\sum_{y\in Y} f\left(y|\mathbb{P}\right)\int_{\phi\in\Phi} \frac{f\left(y|\phi,\mathbb{P}\right)}{f\left(y|\mathbb{P}\right)} f\left(\phi\right)\ln f\left(y|\phi,\mathbb{P}\right)d\phi \\
&\quad +\sum_{y\in Y} f\left(y|\mathbb{P}\right)\int_{\phi\in\Phi} \frac{f\left(y|\phi,\mathbb{P}\right)}{f\left(y|\mathbb{P}\right)} f\left(\phi\right)\ln f\left(y|\mathbb{P}\right)d\phi \\
&\quad -\sum_{y\in Y} f\left(y|\mathbb{P}\right)\int_{\phi\in\Phi} \frac{f\left(y|\phi,\mathbb{P}\right)}{f\left(y|\mathbb{P}\right)} f\left(\phi\right)\ln f\left(\phi\right)d\phi \\
&= \underbrace{-\int_{\phi\in\Phi}\sum_{y\in Y} f\left(y|\phi,\mathbb{P}\right)\ln f\left(y|\phi,\mathbb{P}\right)f\left(\phi\right)d\phi}_{\xi(Y|\Phi,\mathbb{P})} \\
&\quad \underbrace{+\sum_{y\in Y}\underbrace{\int_{\phi\in\Phi} f\left(y|\phi,\mathbb{P}\right)f\left(\phi\right)d\phi}_{f(y|\mathbb{P})}\ln f\left(y|\mathbb{P}\right)}_{-\xi(Y|\mathbb{P})} \\
&\quad \underbrace{-\int_{\phi\in\Phi} f\left(\phi\right)\ln f\left(\phi\right)d\phi}_{\xi(\Phi)}
\end{aligned} \tag{12}
$$

3.2.3 Divergence

Let $\mathcal{N}$ be a constant such that $\int_{\phi\in\Phi}\frac{1}{\mathcal{N}}d\phi = 1$. The Kullback-Leibler divergence for the random variable Φ relative to a uniform distribution can be defined (using the entropy definitions given) as:

$$
Div\left(\Phi|Y,\mathbb{P}\right) = -\xi\left(\Phi|Y,\mathbb{P}\right)+\ln(\mathcal{N})
$$

For a fixed $\mathcal{N}$, maximising expected divergence from the uniform distribution is therefore, equivalent to minimising the entropy of the posterior. For a given prior, then given [10], for high divergence, we need a high value of

$$
\xi\left(Y|\mathbb{P}\right)-\xi\left(Y|\Phi,\mathbb{P}\right) \tag{13}
$$

Note that in general $\xi\left(Y|\mathbb{P}\right)\geq\xi\left(Y|\Phi,\mathbb{P}\right)$.

3.2.4 Choosing an Informative Task

Imagine, we only have one task $\mathbb{P} = \mathcal{P}_1$ such that

$$\begin{aligned}
\xi\left(Y_1|\Phi,\mathcal{P}_1\right) &= -\int_{\phi\in\Phi}\sum_{y=0,1} f\left(y|\phi,\mathcal{P}_1\right)\ln f\left(y|\phi,\mathcal{P}_1\right) f\left(\phi\right)d\phi && (14)\\
\xi\left(Y_1|\mathcal{P}_1\right) &= -\sum_{y\in 0,1} f\left(y|\mathcal{P}_1\right)\ln f\left(y|\mathcal{P}_1\right)\\
& \text{where} && (15)\\
f\left(y|\mathcal{P}_1\right) &= \int_{\phi\in\Phi} f\left(y|\phi,\mathcal{P}_1\right) f\left(\phi\right)d\phi && (16)
\end{aligned}$$

For very large N draws of ϕ_n from the prior $f(\phi)$, we can simulate these quantities as:

$$\begin{aligned}
\hat{\xi}\left(Y_1|\Phi,\mathcal{P}_1\right) &= N^{-1}\sum_{n=1}^{N}\sum_{y=0,1} f\left(y|\phi_n,\mathcal{P}_1\right)\ln f\left(y|\phi_n,\mathcal{P}_1\right)\\
\hat{\xi}\left(Y_1|\mathcal{P}_1\right) &= -\sum_{y=0,1}\hat{f}\left(y|\mathcal{P}_1\right)\ln\hat{f}\left(y|\mathcal{P}_1\right)\\
& \text{where} && (17)\\
\hat{f}\left(y|\mathcal{P}_1\right) &= N^{-1}\sum_{n=1}^{N} f\left(y|\phi_n,\mathcal{P}_1\right) && (18)
\end{aligned}$$

A task is therefore estimated to be "singularly informative" if it has a large value for $\hat{\xi}\left(Y|\mathcal{P}_1\right) - \hat{\xi}\left(Y|\Phi,\mathcal{P}_1\right)$.

3.2.5 Choosing an Informative Pair of Tasks

Now imagine, we have only two tasks $\mathbb{P} = \left(\mathcal{P}_1,\mathcal{P}_2\right)$ such that $Y = (y_1, y_2)$. Let us assume that y_1 has already been chosen as highly informative according to the criteria above. Our task is then to choose y_2. The entropy functions obey the following (under conditional independence $f\left(y_2|y_1,\phi\right) = f\left(y_2|\phi\right)$)

$$\begin{aligned}
\xi\left(Y|\Phi,\mathbb{P}\right) &= \xi\left(Y_1|\Phi,\mathcal{P}_1\right) + \xi\left(Y_2|\Phi,\mathcal{P}_2\right) && (19)\\
\xi\left(Y|\mathbb{P}\right) &= \xi\left(Y_1|\mathcal{P}_1\right) + \xi\left(Y_2|Y_1,\mathbb{P}\right)
\end{aligned}$$

This is follows from Proof B.

Proof B

Conditional Entropy is defined as:

$$\xi\left(y_2|y_1,\mathbb{P}\right)=\sum_{y_2\in 0,1}\ln f\left(y_2|y_1,\mathbb{P}\right)f\left(y_2,y_1\right)$$

Therefore, if we take:

$$\begin{aligned}
\xi\left(Y|\mathbb{P}\right) &= -\sum_{y\in Y}f\left(y_1,y_2|\mathbb{P}\right)\ln f\left(y_1,y_2|\mathbb{P}\right)\\
&= -\sum_{y\in Y}f\left(y_1,y_2|\mathbb{P}\right)\left(\ln f\left(y_2|y_1\mathbb{P}\right)+\ln f\left(y_1\mathbb{P}\right)\right)\\
&= -\sum_{y\in Y}f\left(y_1,y_2|\mathbb{P}\right)\ln f\left(y_2|y_1\mathbb{P}\right)-\sum_{y\in Y}f\left(y_1,y_2|\mathbb{P}\right)\ln f\left(y_1\mathbb{P}\right)\\
&= -\sum_{y\in Y}f\left(y_1,y_2|\mathbb{P}\right)\ln f\left(y_2|y_1\mathbb{P}\right)+\xi\left(y_1|\mathbb{P}\right)\\
&= \xi\left(y_2|y_1,\mathbb{P}\right)+\xi\left(y_1|\mathbb{P}\right)
\end{aligned}$$

If there is conditional independence then it follows that:

$$\sum_{y\in Y}f\left(y_1,y_2|\mathbb{P}\right)\ln f\left(y_2|y_1\mathbb{P}\right)=\sum_{y\in Y}f\left(y_1,y_2|\mathbb{P}\right)\ln f\left(y_2|\mathbb{P}\right)=\xi\left(y_2|\mathbb{P}\right)$$

Therefore, taking $y_1$as fixed, we wish to choose $\mathcal{P}_2$ so as to minimise

$$\xi\left(Y|\Phi,\mathbb{P}\right)-\xi\left(Y|\mathbb{P}\right)=k+\xi\left(Y_2|\Phi,\mathcal{P}_2\right)-\xi\left(Y_2|Y_1,\mathbb{P}\right)$$

We can calculate the quantity $\xi\left(Y_2|\Phi,\mathcal{P}_2\right)$ in exactly the same way as we calculated $\xi\left(Y_1|\Phi,\mathcal{P}_1\right)$. However, $\xi\left(Y_2|Y_1,\mathbb{P}\right)$ needs to be calculated differently to $\xi\left(Y_2|\mathcal{P}_2\right)$. Recalling that

$$\begin{aligned}
\xi\left(Y_2|Y_1,\mathbb{P}\right) &= \sum_{y_2\in 0,1}\sum_{y_1\in 0,1}f\left(y_1,y_2|\mathbb{P}\right)\ln f\left(y_2|y_1,\mathbb{P}\right) \qquad (20)\\
f\left(y_1,y_2|\mathbb{P}\right) &= \int_{\phi\in\Phi}f\left(y_1,y_2|\phi,\mathbb{P}\right)f\left(\phi\right)d\phi\\
f\left(y_2|y_1,\mathbb{P}\right) &= \int_{\phi\in\Phi}f\left(y_1|y_2,\phi,\mathbb{P}\right)f\left(\phi\right)d\phi
\end{aligned}$$

Both the joint and the conditional probabilities can therefore be simulated using N draws of ϕ_n from the prior $f\left(\phi\right)$

$$\begin{aligned}
\hat{f}\left(y_1,y_2|\mathbb{P}\right) &= N^{-1}\sum_{n=1}^{N}f\left(y_1,y_2|\phi_n,\mathbb{P}\right) \qquad (21)\\
\hat{f}\left(y_2|y_1,\mathbb{P}\right) &= \frac{\hat{f}\left(y_1,y_2|\mathbb{P}\right)}{\hat{f}\left(y_1|\mathbb{P}\right)}\text{ where }\hat{f}\left(y_1|\mathbb{P}\right)=N^{-1}\sum_{n=1}^{N}f\left(y_1|\phi_n,\mathcal{P}_1\right)
\end{aligned}$$

Then for any two pairs

$$\hat{\xi}\left(Y_2|Y_1,\mathbb{P}\right)=\sum_{y_2\in 0,1}\sum_{y_1\in 0,1}\hat{f}\left(y_1,y_2|\mathbb{P}\right)\ln\hat{f}\left(\left(y_2|y_1,\mathbb{P}\right)\right)\tag{22}$$

Accordingly this conditional entropy can be calculated for any set of tasks.

3.2.6 Simulating the Full Entropy of a Set of Prospect Pairs

Because the set of all possible outcomes for y is of size 2^T it is unfeasible to sum over this set. However, this quantity can be simulated for any set of draws

$$\xi\left(Y|\Phi,\mathbb{P}\right)=-\int_{\phi\in\Phi}\sum_{y\in Y}f\left(y,\phi|\mathbb{P}\right)\ln f\left(y|\phi,\mathbb{P}\right)d\phi\tag{23}$$

Using this we can therefore:

- Take a draw from $f(\phi)$ then a draw from $f(y|\phi,\mathbb{P})$ to obtain a joint draw (y_s,ϕ_s) and then we can calculate $f(y_n|\phi_s,\mathbb{P})$
- Repeat S times and calculate the average $\hat{\xi}\left(Y|\phi,\mathbb{P}\right)=S^{-1}\sum_{s=1}^{S}\ln f\left(y_s|\phi_s,\mathbb{P}\right)$

Next, since

$$\begin{aligned}\xi\left(Y|\mathbb{P}\right) &= \sum_{y\in Y}f\left(y|\mathbb{P}\right)\ln f\left(y|\mathbb{P}\right) \\ f\left(y|\mathbb{P}\right) &= \int_{\phi\in\Phi}f\left(y|\phi,\mathbb{P}\right)f\left(\phi\right)d\phi\end{aligned}\tag{24}$$

- Take a draw from $f(\phi)$ then a draw from $f(y|\phi)$ to obtain y_s
- Obtain an estimate of $f(y|\mathbb{P})$ as $N^{-1}\sum_{n=1}^{N}f\left(y_s|\phi_n\right)$ where ϕ_n are draws from the prior distribution
- Repeat S times to obtain

$$\hat{\xi}\left(Y\right)=S^{-1}\sum_{s=1}^{S}\ln\left(N^{-1}\sum_{n=1}^{N}f\left(y_s|\phi_n\right)\right)\tag{25}$$

where ϕ_n are draws from $f(\phi)$

3.2.7 Constructing Prospect Sets

A truly efficient design would start with a first prospect pair that was singularly informative, then we could proceed to find pairs that were highly setwise informative. However, this is computationally impractical given the complexity of the calculations due to the very large number of combinations for $(y_T,...,y_1)$, for which there are 2^T possible outcomes. The calculations above would therefore require that $\hat{f}(y_T, ..., y_1)$ be calculated for all 2^T cases which rapidly becomes infeasible. Given the infeasibility of the 'setwise' approach, we adopt a second best solution. The approach used here is to ensure that all prospect pairs are highly 'pairwise informative'. It is relatively straight forward to construct the matrix of all values of the information content of a given prospect pair, along with the additional information of one pair, given another.

Define:

$$\xi_{ii} = -\xi(Y_i|\Phi, \mathbb{P}) \text{ and } \xi_{ij} = -\xi(Y_j|\Phi, \mathbb{P}) + \xi(Y_j|Y_i, \mathbb{P}) \text{ if } i \neq j$$

A procedure to construct the prospect set which we have adopted (having decided on a number of prospect Pairs T) is as follows:

- Make a large number of draws from the prior $f(\phi)$
- Construct a very large number of candidate prospect pairs $\mathbb{P}$ of size T_0. In so doing ensure that initial set does not include tasks that would lead to identical choices across the entire range of preferences and contains no repeated tasks
- Simulate all ξ_{ij} for all prospect pairs (given the priors $f(\phi)$)
- Rank them such that $\xi_{11} > \xi_{22} > ... > \xi_{LL}$ (i.e. in terms of their singular informativeness)
- For a given (small) κ eliminate prospect pairs where $\xi_{kj} < \kappa$ for any i and $j > i$ to leave a set of size M
- Allow κ increase to the point where $M = T$ (the desired number of prospect pairs)

This procedure does not promise to maximise divergence, but it does ensure that each pair of pairs is non-redundant in an informational sense (i.e. they are pairwise informative). That is, T prospect pairs will have been constructed where the one with the highest information is included and any other prospect that has been included has a minimal pairwise information contribution of κ, where κ is made as high as possible.

4 Demonstration

In this section, we give a brief illustration of the comparative efficiency of designs using the method described in this paper. The design is compared with two other lottery sets of the same dimension (i.e. 100 lotteries) given in Appendix A of Harrison and Swarthout (2016). These design draw heavily on designs introduced and used by Loomes and Sugden (1998). In each case the simulation of preferences was done using 10,000 draws of preferences from the same "priors" outlined in the experimental design section. We have generated three sets of lotteries and these are given in Excel spreadsheet (Simulation Lotteries) that accompanies this paper.

The growth in the efficiency/entropy measure (in 23) at each point in the lottery set is shown in Figure 1.

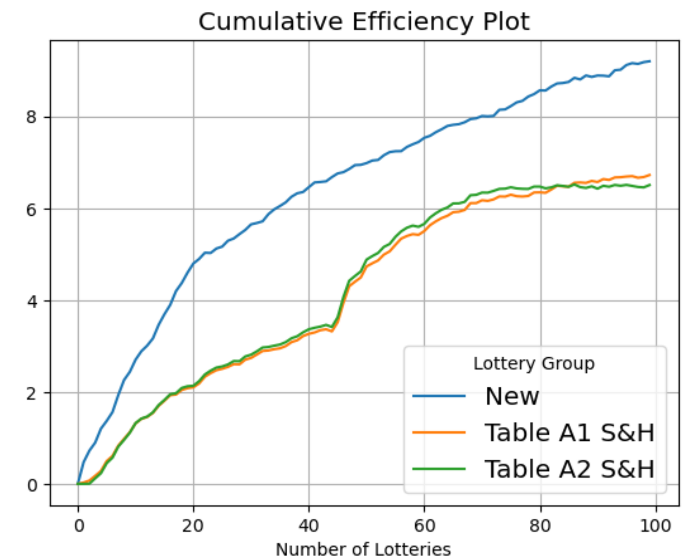

Figure 1: Cumulative Efficiency Plot of New Design Compared to Harrison and Swarthout (2016)

:

In Figure 1, the two existing lottery sets show a "staggered" growth path since the calculation is at each point in the ordered set where the increase in the rate of efficiency gain is occurs in the middle of the lottery sets as this is when loss/mixed lotteries enter the calculation. The new set of lotteries we have generated (where we have in each case a sure thing) is not ordered in this way hence the smoother nature of the growth path in efficiency gain. As can be seen by the results shown in Figure 1, for every given number of lotteries the aggregate efficiency measure is always higher than for the corresponding lotteries in Harrison and Swarthout (2016). This difference in the design of lotteries implies that the level of information that will be revealed is always greater for the set of lotteries we have generated. The potential benefit of the efficiency gain are that fewer lotteries are required to reveal the same level of information about risk preferences.

5 Concluding Comments

In this we note, we explain how to design and implement an experimental approach capable of accurately characterizing people's risk preferences, assuming that they behaved in accordance with Prospect Theory (PT). To accomplish this, we have taken as our guiding principle that the tasks assigned to participants are informative, meaning that individuals with different preferences would make distinct choices. Given the flexibility of PT models, the heterogeneous preferences of people cannot be captured by asking them to perform only a few tasks. Yet, simply increasing the number of tasks that people must complete, may not improve the potential to capture their risk preferences should some choices establish preferences already revealed by previous choices. Therefore, we introduced an approach to designing a survey that reduces the propensity to include noninformative or informationally redundant tasks. Our research adds to a literature that has considered how tasks can be designed efficiently and implemented effectively, such as Adaptive Design Optimization (ADO). Although the approach we have taken is different to those previously employed in the literature, we are able to generate a set of tasks that are informative without imposing strong prior beliefs about model parameters. We also note, that the approach that we have described will not yield an "efficient" set of lottery pairs but it does extend aspects of lottery design that warrant further investigation.

References

Amador-Hidalgo, L., Brañas-Garza, P., Espín Martín, A.M., García Muñoz, T.M. and Hernández, A. (2021). Cognitive abilities and risk-taking: Errors, not preferences. **European Economic Review**, 134 may, 103694.

Andersson, O., Holm, H.J., Tyran, J.R. and Wengström, E. (2016). Risk aversion relates to cognitive ability: Preferences or noise?. **Journal of the European Economic Association**, 14(5): 1129-1154.

Andersen, S., Cox, J.C., Harrison, G.W., Lau, M.I., Rutström, E.E. and Sadiraj, V. (2018). Asset integration and attitudes toward risk: Theory and evidence. **Review of Economics and Statistics**, 100(5): 816-830.

Balcombe, K.G. and Fraser, I.M. (2015). Parametric preference functionals under risk in the gain domain: A Bayesian analysis. **Journal of Risk and Uncertainty**, 50: 161-187.

Balcombe, K.G., Bardsley, N., Dadzie, S.and Fraser, I.M. (2019). Estimating parametric loss aversion with prospect theory: Recognising and dealing with size dependence. **Journal of Economic Behavior and Organization**, 162: 106-119.

Bauer, M., Chytilova, J. and Morduch, J. (2012). Behavioral Foundations of Microcredit: Experimental and Survey Evidence from Rural India. **American Economic Review**, 102:1118–39.

Bernheim, B.D. and Sprenger, C. (2020). On the Empirical Validity of Cumulative Prospect Theory: Experimental Evidence of Rank-Independent Probability Weighting. **Econometrica**, 88: 1363-1409.

Bernheim, B.D., Royer, R. and Sprenger, C. (2022). Robustness of Rank Independence in Risky Choice. **American Economic Review**, 112: 415-420.

Binswanger, H. (1980). Attitudes toward Risk: Experimental Measurement in Rural India. **American Journal Agricultural Economics**. 62 (3): 395–407.

Booij, A. S., van Praag, M.S. and van de Kuilen, G. (2010). A Parametric Analysis of Prospect Theory's Functionals for the General Population. **Theory and Decision**, 68: 115-148.

Bougherara, D., Gassmann, X., Piet, L. and Reynaud, A. (2017). Structural estimation of farmers' risk and ambiguity preferences: a field experiment. **European Review of Agricultural Economics**, 44(5): 782-808.

Broomell, S. B. and Bhatia, S. (2014). Parameter recovery for decision modeling using choice data. **Decision**, 1(4): 252-274.

Bruhin, A., Fehr-Duda, H. and Epper, T. (2010). Risk and Rationality: Uncovering Heterogeneity in Probability Distortion. **Econometrica**, 78: 1375-1412.

Buschena, D.E and Zilberman, D. (1999). Testing the Effects of Similarity on Risky Choice: Implications for Violations of Expected Utility. **Theory and Decision,** 46: 251–276.

Buschena, D.E.and Atwood, J.A. (2011). Evaluation of similarity models for expected utility violations. **Journal of Econometrics,** 162: 105–113.

Cavagnaro, D.R., Gonzalez, R., Myung, J.I. and Pitt, M.A. (2013a). Optimal decision stimuli for risky choice experiments: An adaptive approach. **Management Science**, 59(2): 358-375.

Cavagnaro, D.R., Pitt, M.A., Gonzalez, R. and Myung, J.I. (2013b). Discriminating among probability weighting functions using adaptive design optimization. **Journal of Risk and Uncertainty**, 47: 255-289

Charness, G., Gneezy, U. and Imas, A. (2013). Experiential methods: Eliciting risk preferences. **Journal of Economic Behavior & Organization**, 87: 43–51.

Chowdhury, S., Sutter, M. and Zimmermann, K.F. (2022). Economic preferences across generations and family clusters: A large-scale experiment in a developing country. **Journal of Political Economy**, 130(9): 2361-2410.

Crosetto, P. and Filippin, A. (2016). A theoretical and experimental appraisal of four risk elicitation methods. **Experimental Economics**, 19: 613–641.

Drichoutis, A.C. and Lusk, J.L. (2016). What can multiple price lists really tell us about risk preferences? **Journal of Risk and Uncertainty**, 53:89–106.

Eckel, C.C. and Grossman, P.J. (2002). Sex differences and statistical stereotyping in attitudes toward financial risk. **Evolution and Human Behavior**, 23(4): 281-295.

Eckel, C.C. and Grossman, P.J. (2008). Men, women and risk aversion: Experimental evidence. **Handbook of experimental economics results**, 1, pp.1061-1073.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D. and Sunde, U. (2018). Global evidence on economic preferences. **The Quarterly Journal of Economics**, 133(4): 1645-1692.

Freeman, D.J. and Mayraz, G. (2019). Why choice lists increase risk taking. **Experimental Economics**, 22:131–154.

Frydman, C. and Jin, L.J. (2022). Efficient coding and risky choice. **The Quarterly Journal of Economics**, 137(1): 161-213.

Grether, D.M. and Plott, C.R. (1979). Economic Theory of Choice and the Preference Reversal Phenomenon. **American Economic Review**, 69(4): 623-638.

Harrison, G.W. and Ng, J.M. (2016). Evaluating the expected welfare gain from insurance. **Journal of Risk and Insurance**, 83(1): 91-120.

Harrison G.W. and J. Swarthout T (2016). Cumulative Prospect Theory in the Laboratory: A Reconsideration, Experimental Economics Center Working Paper Series 2016-04, Experimental Economics Center, Andrew Young School of Policy Studies, Georgia State University.

Hey, J.D. (2001). Does Repetition Improve Consistency? **Experimental Economics**, 4: 5–54.

Hey, J.D. and Orme, C. (1994). Investigating Generalisations of Expected Utility Theory Using Experimental Data. **Econometrica**, 62(6): 1291-1326.

Holt, C. and Laury, S. (2002). Risk aversion and incentive effects. **American Economic Review**, 92(5): 1644–1655.

Holzmeister, F. and Stefan, M. (2021). The risk elicitation puzzle revisited: Across-methods (in) consistency?. **Experimental Economics**, 24(2): 593-616.

Johnston, R.J., Boyle, K.J., Adamowicz, W., Bennett, J., Brouwer, R., Cameron, T.A., Hanemann, W.M., Hanley, N., Ryan, M., Scarpa, R., Tourangeau, R. and Vossler, C.A. (2017). Contemporary Guidance for Stated Preference Studies. **Journal of the Association of Environmental and Resource Economists**, 4(2): 319-405.

Kahneman, D. and Tversky, A. (1986). Rational choice and the framing of decisions. **Journal of Business**, 59(4): 251-278.

Kpegli, Y.T., Corgnet, B. and Zylbersztejn, A. (2023). All at once! A comprehensive and tractable semi-parametric method to elicit prospect theory components. **Journal of Mathematical Economics**, 104, p.102790.

l'Haridon, O. and Vieider, F.M. (2019). All over the map: A worldwide comparison of risk preferences. **Quantitative Economics**, 10(1): 185–215.

Lichtenstein, S. and Slovic, P. (1971). Reversal of Preferences Between Bids and Choices in Gambling Decisions. **Journal of Experimental Psychology**, 89: 46-55.

Liu, E.M. and Huang, J. (2013). Risk Preferences and Pesticide Use by Cotton Farmers in China. **Journal of Development Economics**, 103: 202-215.

Loomes,G. and Pogrebna, G. (2017). Do Preference Reversals Disappear When We Allow for Probabilistic Choice? **Management Science,** 63(1):166-184.

Loomes, G. and Sugden, R. (1998). Testing Different Stochastic Specifications of Risky Choice. **Economica**, 65: 581-598.

Moffatt, P. G. (2007). Optimal Experimental Design in Models of Decision and Choice. Chapter 15: 357-375, In Boumans M. (ed.) **Measurment in Economics: A Handbook**, Academic Press: London.

Moffatt, P.G. (2015). **Experimetrics: Econometrics for Experimental Economics**. Palgrave Macmillan, London.

Müller, W.G. and Ponce de Leon, A.C.M. (1996). Optimal Design of an Experiment in Economics, **The Economic Journal**, 106 (434): 122–127.

Murphy, R.O. and W. ten Brincke, R.H. (2018). Hierarchical Maximum Likelihood Parameter Estimation for Cumulative Prospect Theory: Improving the Reliability of Individual Risk Parameter Estimates. **Management Science**, 64(1): 308-326.

Nilsson, H., Rieskamp, J. and Wagenmakers, E.J. (2011). Hierarchical Bayesian parameter estimation for cumulative prospect theory, **Journal of Mathematical Psychology**, 55(1): 84-93.

Olschewski, S., Sirotkin, P. and Rieskamp, J. (2021). Empirical underidentification in estimating random utility models: The role of choice sets and standardizations. **British Journal of Mathematical and Statistical Psychology**, 75: 252-292.

Pedroni, A., Frey, R., Bruhin, A., Dutilh, G., Hertwig, R. and Rieskamp, J. (2017). The risk elicitation puzzle. **Nature Human Behaviour**, 1(11): 803-809.

Pfeiffer, J., Meißner, M., Brandstätter, E., Riedl, R., Decker, R. and Rothlauf, F. (2014). On the influence of context-based complexity on information search patterns: An individual perspective. **Journal of Neuroscience, Psychology, and Economics**, 7(2): 103–124.

Regier, D.A., Watson, V., Burnett, H. and Ungar, W.J. (2014). Task Complexity and Response Certainty in Discrete Choice Experiments: An Application to Drug Treatments for Juvenile Idiopathic Arthritis. **Journal of Behavioral and Experimental Economics**, 50: 40-49.

Rieskamp, J. (2008). The Probabilistic Nature of Preferential Choice. **Journal of Experimental Psychology: Learning, Memory, and Cognition**, 34(6): 1446–1465.

Sloman, S.J., Cavagnaro, D. and Broomell, S.B. (2023). Knowing what to know: Implications of the choice of prior distribution on the behavior of adaptive design optimization. arXiv preprint arXiv:2303.12683.

Stott, H.P. (2006). Cumulative prospect theory's functional menagerie. **Journal of Risk and Uncertainty**, 32: 101–130.

Tanaka, T., Camerer, C.F. and Nguyen, Q. (2010). Risk And Time Preferences: Linking Experimental And Household Survey Data From Vietnam. **American Economic Review**, 100(1): 557-571.

Tversky, A. and Kahneman, D. (1981). The framing of decisions and the psychology of choice. **Science**, 211: 453-58.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. **Journal of Risk and Uncertainty,** 5(4): 297–323.

Wakker, P.P. (2010). **Prospect Theory: For Risk and Ambiguity**, Cambridge University Press.

Ward, P.S., and Singh, V. (2015). Using Field Experiments to Elicit Risk and Ambiguity Preferences: Behavioural Factots and the Adoption of New Technologies in Rural India. **Journal of Development Studies**, 51(6): 707-724.

Yang, J., Pitt, M.A., Ahn, W.Y. and Myung, J.I. (2021). ADOpy: a python package for adaptive design optimization. **Behavior Research Methods**, 53: 874-897.

Zilker, V. and Pachur, T. (2021). Nonlinear probability weighting can reflect attentional biases in sequential sampling. **Psychological Review,** (Early View).