



ORIGINAL PAPER

The Multivariate Heterogeneous Preference estimator for Switching Multiple Price List data

Anna Conte¹ , Peter G. Moffatt²  and Mary Riddell³ 

¹Department of Statistical Sciences, Sapienza University of Rome, Rome, Italy

²School of Economics, University of East Anglia, Norwich, UK

³Department of Economics, University of Nevada, Las Vegas, USA

Corresponding author: Anna Conte; Email: anna.conte@uniroma1.it

(Received 6 September 2021; revised 12 April 2023; accepted 4 March 2026)

Abstract

The Multiple Price List (MPL) and Switching Multiple Price List (sMPL) provide a useful framework for estimating preference parameters, most usually risk aversion, from a sample of experimental subjects or survey respondents. In this paper, we consider designs in which *more than one* sMPL is presented to each subject, allowing *more than one* preference parameter to be estimated simultaneously, and we propose a consistent estimator in this setting – the Multivariate Heterogeneous Preference (MHP) estimator. Focusing on the bivariate case of two sMPLs and two preference parameters, we demonstrate that non-standard econometric techniques, namely Monte Carlo integration with importance sampling, are required to implement the MHP estimator. Via a Monte Carlo exercise, we show that our estimator has good finite-sample properties. Finally, we apply the MHP estimator to a real data set and compare the estimates to those obtained using an inconsistent estimator applied in previous studies.

Keywords: importance sampling; Monte Carlo integration; risk preference

JEL Classification: C51; C52; C91; D81

1. Introduction

The Multiple Price List (MPL) has become a very popular method in Experimental Economics for eliciting individuals' preferences. Early applications of the method included elicitation of willingness to pay for a good (Kahneman et al., 1990), measurement of risk attitudes (Binswanger, 1980), and elicitation of individual discount rates (Coller and Williams, 1999).¹ For the willingness-to-pay application, subjects see an ordered list of prices, and are asked to respond with “yes” or “no” to each

¹The number of articles using MPL to refer to is enormous, and reviewing them is outside the main scope of this work. Recent examples include Bellemare and Shearer (2010), Bruhin et al. (2010), Garbarino et al. (2010), Abdellaoui et al. (2011), Brañas-Garza and Rustichini (2011), von Gaudecker et al. (2011), Voors et al. (2012), Bosch-Domènech and Silvestre (2013), Tanaka and Munro (2013) Andersen et al. (2014), Aycinena et al. (2014), Ebert and Wiesen (2014), Ziegelmeier and Ziegelmeier (2014), Astebro et al. (2014), Drichoutis and Nayga (2015), Bolton et al. (2015), Santos-Pinto et al. (2015), Sprenger (2015), Crosetto and Filippin (2016), Csermely and Rabas (2016), Drichoutis and Lusk (2016), Ferecatu and Öncüler (2016), Bruner (2017), Habib et al. (2017), Alonso et al. (2018), Attanasi et al. (2018), Charness et al. (2018), He and Hong (2018), Vieider (2018), Bradford et al. (2019), Eijkelenboom et al. (2019), Freeman et al. (2019), Parslow et al. (2019), Vieider et al. (2019) Andersson et al. (2020), Beauchamp et al. (2020), Charness et al. (2020), Cheung (2020), Bougherara et al. (2021), Wegmann et al. (2021), Yu et al. (2021), Bonjean (2023), Bougherara et al. (2017), Brañas-Garza et al. (2022), Martínez et al. (2023),

price. For the other two applications (risk attitude and discount rate), subjects see an ordered list of dichotomous choice tasks, and are asked to indicate their choice in each. A particularly popular MPL designed for the purpose of eliciting risk attitude is that of [Holt and Laury \(2002, HL henceforth\)](#).

The Switching Multiple Price List (sMPL) tweaks the basic design of the MPL by asking subjects to state in which row of the MPL table they would like to switch from one column to the other. The first implementation of this design was by [Gonzalez and Wu \(1999\)](#). As noted by [Andersen et al. \(2006\)](#), the sMPL has the obvious advantage that it induces monotonicity of preferences (since only one switch-point is permitted) and an additional advantage over lottery-choice and certainty-equivalent methods in that it is comparatively simple for subjects to master. Early users of the sMPL design were [Eckel and Grossman \(2002\)](#) and [Tanaka et al. \(2010\)](#), and these studies have become reference points for many other laboratory and field experiments in economics.² The popularity of such a design is also encroaching upon other fields of economics, including agricultural, development, energy and environmental and resource economics, and even medicine.³

Despite the overwhelming popularity of MPLs and their variants, some researchers have questioned their incentive compatibility, notably [Brown and Healy \(2018\)](#). Here, an important distinction needs to be made between theoretical and behavioural incentive compatibility ([Kaas and Ruprecht, 2006; Danz et al., 2022; Danz et al., 2024](#)). It can easily be shown that, given the standard practice of selecting one of the rows of the list at random and playing the selected lottery for real, and given the auxiliary assumption that decision-makers maximise Expected Utility, the sMPL is theoretically incentive-compatible in the same way that popular incentive schemes such as Random Lottery Incentive (RLI) and Becker-deGroot-Marschak (BDM) are theoretically incentive compatible. We acknowledge that sMPL may not be behaviourally incentive compatible, especially in the light of [Brown and Healy's](#) findings (and there is a substantial literature questioning the behavioural incentive compatibility of other popular schemes such as RLI and BDM, see [Bardsley et al., 2010, Chapter 6](#)). However, it must also be noted that [Brown and Healy's](#) results actually relate to MPL and not sMPL. We are not aware of any work that brings into question the behavioural incentive compatibility of sMPL.

Data collected via the sMPL may be used to estimate the preference parameter of interest. In the context of the risk-attitude application, given an assumed parametric utility function, the subject's reported switch-point implies a range of values of the risk-aversion parameter for that subject. The appropriate estimation approach is the well-established technique of interval regression,⁴ and the estimation results provide estimates of the distributional parameters of risk aversion over the population, as well as treatment effects and also estimates of the effects of individual characteristics on risk aversion. Since this model is built on the idea that each individual possesses their own risk-aversion parameter, and this parameter varies continuously over the population, we will refer to this model as the Heterogeneous Preference (HP) model.

While the HP model has attractive features, including ease of estimation, it also has one undesirably restrictive feature: it only allows estimation of the distribution of a single preference parameter.⁵ A closely related problem is that an auxiliary assumption – typically expected utility (EU) maximisation – must be made in order to obtain the intervals of risk-aversion implied by each

[Mograbi \(2022\)](#), [Meissner et al. \(2023\)](#). A list of articles using MPL to elicit willingness to pay in laboratory and lab-in-the-field experiments, and sample surveys can be found in [Jack et al. \(2022\)](#).

²See, among others, [Loomes et al. \(2003\)](#); [Loomes et al. \(2010\)](#), [Riddell \(2012\)](#), [Riddell and Kolstoe \(2013\)](#), [Sugden et al. \(2013\)](#), [Isoni et al. \(2016\)](#), [He and Villeval \(2017\)](#), [Rodriguez-Lara and Ponti \(2017\)](#), [Boyd and Bellemare \(2022\)](#).

³See, among others, [Nguyen \(2010\)](#); [Nguyen \(2011\)](#), [Reynaud and Couture \(2012\)](#), [Bocquého et al. \(2013\)](#), [Liu \(2013\)](#), [Liu and Huang \(2013\)](#), [Menapace et al. \(2013\)](#), [Liebenehm and Waibel \(2014\)](#), [Menapace et al. \(2016\)](#), [Bartczak et al. \(2017\)](#), [Goldzahl \(2017\)](#), [Vollmer et al. \(2018\)](#), [Zhu et al. \(2020\)](#), [Kreft et al. \(2021\)](#), [Bonan et al. \(2022\)](#), [Shin et al. \(2022\)](#).

⁴Interval regression is now available in most econometric packages.

⁵This point is stressed by [Harrison and Rutström \(2008, p. 67\)](#): "limitation of this approach [...] is that it restricts the analyst to utility functions that can characterize risk attitudes using one parameter".

switch-point. A huge literature has established the tendency for subjects to deviate from EU (see, e.g., Starmer, 2000), and the deviation from EU is often modelled parametrically by assuming a probability weighting function. An obvious question that follows is whether it is possible to extend the sMPL framework in order to allow for non-EU preferences and how to estimate the parameters of the assumed probability weighting functions.

This becomes possible when there is more than one sMPL (see Section 3). Stated simply, if there are S sMPL tasks per subject and K preference parameters, and $S \geq K$, it is possible to estimate the parameters of the K -variate distribution of these preference parameters. If $S > K$ it is also possible to estimate a parameter representing within-subject variation, with $S - K$ interpreted as the number of degrees-of-freedom (see also Appendix C). For reasons that will become clear, we will be focusing on the case $S = K$, in which there are no degrees-of-freedom available for identifying within-subject variation. The model used to estimate the joint distribution of the preference parameters will be referred to as the Multivariate Heterogeneous Preference (MHP) model. In fact, we will further restrict attention to the case, $S = K = 2$.

Designs with multiple sMPLs have already been used by Tanaka et al. (2010), Riddell (2012) and Riddell and Kolstoe (2013), among others. All of these designs in principle allow simultaneous estimation of both risk-aversion and probability weighting. However, these researchers used econometric tools that cannot claim to produce consistent estimates. Tanaka et al. (2010) applied linear regression techniques to the midpoints of the implied intervals to estimate the curvature of the utility function.⁶ Liu (2013) also uses linear regression on midpoints to separately estimate value-function curvature, loss aversion and probability weighting coefficient. Apart from the obvious problem of open intervals not having midpoints, the inconsistency (in estimation) of such an approach has been established by Stewart (1983). Riddell and Kolstoe (2013) applied interval regression to their sMPL data. However, they estimated risk aversion and probability weighting separately, hence assuming independence between the two. We are not aware of any studies that have used multiple sMPL data to estimate correlations between preference parameters.

The purpose of this paper is firstly to establish the MHP model as an econometric framework that provides consistent estimates of the parameters of the joint distribution of more than one preference parameter, when the data comes from a design consisting of more than one sMPL. We will then apply the MHP model to data from two sMPL tasks per subject, to estimate the bivariate (between-subject) distribution (including the correlation coefficient) of the two preference parameters, risk aversion and probability weighting. As explained above, since the number of parameters in the joint distribution is equal to the number of sMPLs, it is necessary to assume away within-subject variation, that is, to assume that all decisions, for a given subject, are made using the same pair of preference parameters and without error.

In the case of two sMPLs, the MHP model may be viewed as a bivariate generalisation of the interval regression model, and we might describe the data as “bivariate interval data”. However, the extension to two dimensions brings about a non-standard feature: the set of combinations of the two parameters that is implied by any pair of switch-points is non-rectangular; in fact, it is not even the case that the set is bounded by straight lines. The region implied by each switch-point combination may be described as a curvilinear quadrilateral. For consistent estimation of the distributional parameters, the contributions to the likelihood function must be computed as the volume under the assumed bivariate density function over the curvilinear quadrilateral indicated by the observed pair of switch-points. Monte Carlo integration (Hammersley and Handscomb, 1964) is required for this purpose.

The econometric problems do not end there. A further problem is that some of the regions may be too small and/or too far from the centre of the estimated distribution for the required probability to be estimated with any accuracy, given a realistic number of simulations. This is known as the “rare-event

⁶Tanaka et al. (2010) also use interval regression to estimate a loss-aversion parameter.

sampling” problem (see, among others, Rubinstein and Kroese, 2017). For this reason, the importance sampling technique (Gouriéroux and Monfort, 1996) is required.

Hence it is clear that consistent estimation in the multiple sMPL setting is – in contrast to the single sMPL case – only possible using fairly advanced econometric techniques. To our knowledge, this problem has not previously been solved, and this paper aims to meet this econometric challenge by developing a consistent estimator for the bivariate case. We shall also investigate how the results from our consistent estimation procedure differ from those obtained from computationally simpler methods such as linear regression using midpoints.

The remainder of the paper is organised as follows. Section 2 presents a popular MPL and explains how the HP model can be applied to this design when preferences are characterised by one single parameter. The drawbacks of the MPL are highlighted and it is explained how using the sMPL overcomes these drawbacks. Section 3 provides an example of a design consisting of two sMPLs, and explains how a subject’s pair of switch-points from two sMPLs can be used to deduce a set of parameter combinations in which that subject’s parameter values are known to lie. Section 4 describes the estimation approach based on midpoints, adopted by some other researchers, and stresses its drawbacks. Section 5 develops the MHP estimator using Monte Carlo Integration with Importance Sampling. In Section 6, we report the MHP estimation results from a real data set, and compare them with those obtained from the midpoint estimator. In Section 7, we suggest some extensions of the approach, we provide a set of recommendations for practitioners needing to choose between different approaches, and we also highlight an important caveat. Section 8 concludes.

2. The basic MPL design and the Heterogeneous Preference model

In this section, we illustrate the use of MPL and sMPL with a single-parameter utility function and show how intervals of values for the relevant parameter are inferred from a subject’s choices. We then discuss how errors in choices can be taken into account when sMPLs are used and also suggest an extension of the interval regression approach that could allow for such errors.

2.1. MPL, sMPL and parameter intervals

As mentioned in Section 1, a very popular example of the Multiple Price List (MPL), in the context of measuring risk aversion, is that of Holt and Laury (2002). We will use this design as an example here, in order to fix some very important ideas. The Holt-Laury MPL is presented in Table 1. It consists of ten rows, each displaying two alternative two-outcome lotteries, A and B. Subjects are asked to indicate their preferred lottery in each row.

Given the outcomes appearing in the lotteries, it is clear that A is the safer lottery and B is the riskier. While the outcomes are fixed between rows, the probabilities change in such a way that option B becomes more attractive as we move down the list. Under the assumption that subjects are Expected Utility (EU) maximisers, and given a choice of a one-parameter utility function, say the popular CRRA utility function $U(x) = x^{1-r}/(1-r)$ where r is the coefficient of relative risk aversion, a subject’s decision in each row determines a range for r as displayed in Figure 1, panel (a).⁷

Subjects with monotonic preferences are expected to choose option A (the safer lottery) for the first few rows and then to switch to lottery B, and to continue choosing B in all subsequent rows. We will refer to the row where the subject switches from one column to the other as the “switch-point”. Each switch-point implies a range of possible values of the parameter of the chosen utility function for that subject. These intervals are shown in the final column of Table 1. Figure 1 provides a diagrammatic explanation of the intervals and related concepts. The interval for a subject who switches in row 6 is shown in Figure 1, panel (b). Such a switching decision implies that lottery A is preferred to lottery B

⁷In this one-parameter utility function, r is the relative risk aversion parameter, taking negative (positive) values for risk-lovers (-averters) and equalling zero for risk-neutrals.

Table 1. The Holt and Laury (2002) design, with ranges of risk aversion parameter implied by each switch-point

Row	Option A	Option B	Expected payoff difference	Range of relative risk aversion
1	1/10 of \$2, 9/10 of \$1.60	1/10 of \$3.85, 9/10 of \$0.10	\$1.17	$r < -1.72$
2	2/10 of \$2, 8/10 of \$1.60	2/10 of \$3.85, 8/10 of \$0.10	\$0.83	$-1.72 < r < -0.95$
3	3/10 of \$2, 7/10 of \$1.60	3/10 of \$3.85, 7/10 of \$0.10	\$0.50	$-0.95 < r < -0.49$
4	4/10 of \$2, 6/10 of \$1.60	4/10 of \$3.85, 6/10 of \$0.10	\$0.16	$-0.49 < r < -0.15$
5	5/10 of \$2, 5/10 of \$1.60	5/10 of \$3.85, 5/10 of \$0.10	-\$0.18	$-0.15 < r < 0.15$
6	6/10 of \$2, 4/10 of \$1.60	6/10 of \$3.85, 4/10 of \$0.10	-\$0.51	$0.15 < r < 0.41$
7	7/10 of \$2, 3/10 of \$1.60	7/10 of \$3.85, 3/10 of \$0.10	-\$0.85	$0.41 < r < 0.68$
8	8/10 of \$2, 2/10 of \$1.60	8/10 of \$3.85, 2/10 of \$0.10	-\$1.18	$0.68 < r < 0.97$
9	9/10 of \$2, 1/10 of \$1.60	9/10 of \$3.85, 1/10 of \$0.10	-\$1.52	$0.97 < r < 1.37$
10	10/10 of \$2, 0/10 of \$1.60	10/10 of \$3.85, 0/10 of \$0.10	-\$1.85	$r > 1.37$

Note: The last column reports the ranges of the relative risk aversion parameter r implied by each switch-point, assuming the CRRA utility function $U(x) = x^{1-r}/(1-r)$.

in the rows 1-5 and that lottery B is preferred instead in rows 6-10. This determines a unique interval of values for r , which is delimited by the two blue vertical lines and contains the only values for r that can explain all ten of these “choices” taken together.⁸

Given the switch-points observed from a sample of experimental subjects, the distribution of r over the population can be estimated using interval regression, with the intervals specified in the final column of Table 1. This is the (univariate) Heterogeneous Preference (HP) model introduced in Section 1.⁹

At first sight, the HP model appears convenient and easy in every respect: MPL choice tasks are straightforward, quick and easy to implement; an interval of values for the parameter of interest can be readily derived from a subject’s choices; the econometric tool to estimate the data (interval regression) is available in the most popular statistical packages.¹⁰

Nonetheless, there are at least two significant problems with the application of the HP model to MPL data, which invalidate it as an approach to the modelling of behaviour under risk. First of all, an unavoidable feature of MPL data is that some subjects switch more than once when scrolling through the list, implying (under the assumption of deterministic revealed preference) non-monotonicity of preferences. Figure 1, panel (c), shows an example of how the values of r determined by the subject’s choice pattern would appear in the case of multiple switch-points, assuming that its real parameter value is between the two vertical lines. HL, for example, report that roughly 10% of subjects display such inconsistencies. Given this, it is perhaps tempting to treat the sequence of choices as a set of (conditionally) independent observations, and to assume some form of within-subject variation in order to explain the multiple switching. However, Brown and Healy (2018) compare two treatments: one in which experimental subjects face 20 decision tasks presented in order and displayed on a single screen, like in the HL design; and another in which the the 20 tasks are presented on separate screens and in random order, with each subject seeing the tasks in a different order. They report that, in the first case, 5% of the subjects exhibit multiple switching, while, in the second case, this percentage rises

⁸For further detail on finding parameter-ranges from switch-points, see Appendix A.1.

⁹Further detail on the use of interval regression to estimate the HP model is provided in Appendix A.1.

¹⁰Ready-made interval regression routines can be directly applied whenever the distribution of the preference parameter is assumed to be normal or lognormal. If the researcher wishes to assume a different distribution, programming the likelihood function is relatively straightforward.

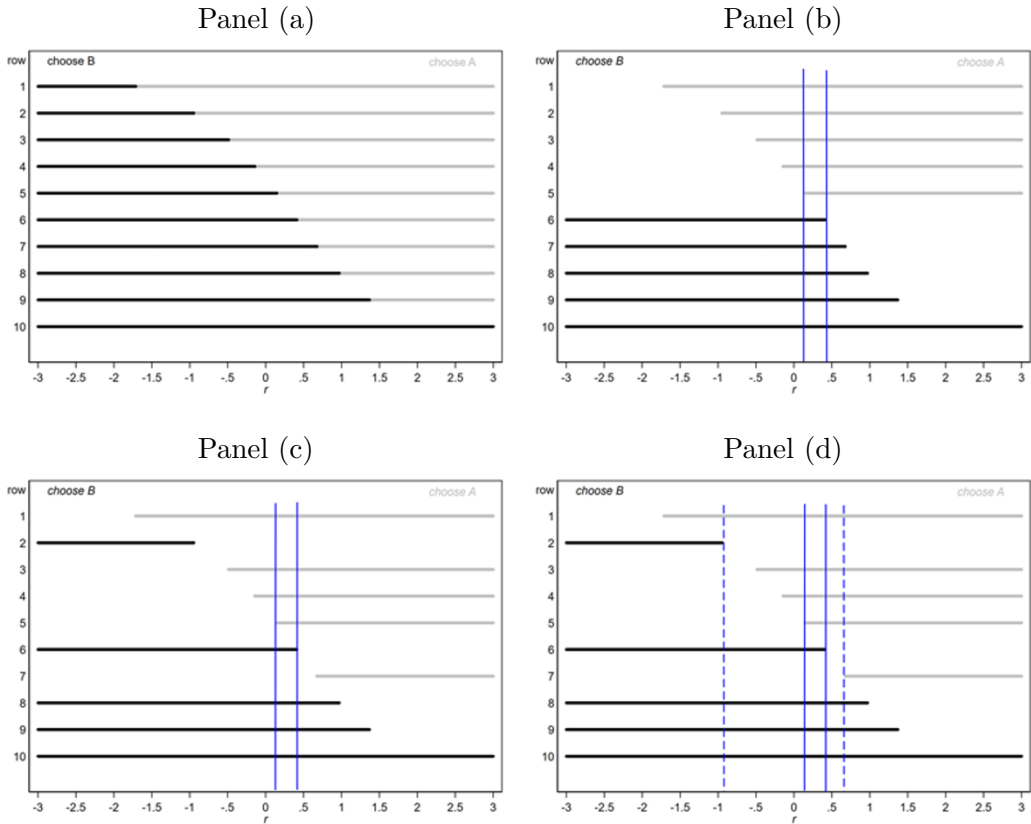


Figure 1. Ranges for r for each row (panel (a)), a single switch in row 6 (panel (b)) and multiple switch-points (bottom panels)

Note: In panel (a), the ranges of r (assuming EU maximisation) for the choice of preferred lottery in each row of Table 1 are displayed, black (grey) if lottery A (B) is chosen. Panel (b) shows the ranges of r which are implied by having switched from lottery A to lottery B in row 6. The vertical lines indicate the range of values for r that are implied by such a decision. Panels (c) and (d) depict the pattern of r values implied by multiple switching behaviour, assuming that the real r is between the two solid vertical lines. In panel (d), the dashed blue lines represent the enlarged interval which is implied by multiple switching. In all figures, left and right bounds are infinite. We note that there are no plausible values of r that could explain a choice of lottery A in row 10, this being strictly dominated by lottery B.

to 32.8% (χ^2 test: p -value = 0.00013).¹¹ The implicit assumption when subjects are presented with choice tasks one-at-a-time in a random order is that the decisions are (conditionally) independent and it is safe to treat them as such. However, the dramatically lower incidence of multiple switching that occurs when the tasks are presented together and in order, as in Table 1, clearly points to subjects having been nudged towards a set of choices with a single switch-point, resulting in a dependence structure between the individual choices which undermines the (conditional) independence hypothesis which is the basis of the majority of behavioural models that are commonly applied to risky choice data. These behavioural models are discussed in the next sub-section.

A solution to the multiple-switching problem is to adopt the switching-variant of the MPL, that is, the sMPL (see, e.g. Andersen et al. 2006). The sMPL approach alters the basic design of an MPL series by asking the subject to report the switch-point directly. Thus, this variant of the MPL approach forces

¹¹Eckel et al. (2004), p. 248, who use a MPL approach but present the choice problems one at a time in a random order, write: “We now believe that scrambling is a bad idea because it results in greater inconsistency and variance of responses.”

preferences to be monotonic.¹² With a single switch-point per subject, the above-mentioned interval regression model can be safely applied to the data.

The second major problem with the MPL (and also with the sMPL) is that it only allows estimation of the distribution of a single preference parameter. If this parameter is risk aversion, an auxiliary assumption such as EU maximisation is required. But it is now recognised beyond any doubt that the EU hypothesis cannot accommodate the vast majority of people's choices under risk (see, e.g. Rabin 1998, Starmer 2000, Bruhin et al. 2010, Conte et al. 2011). Generalisations of EU are required, such as Cumulative Prospect theory (Tversky and Kahneman 1992) or Rank Dependent Utility (RDU) (Quiggin 1982). These models allow for violations of the independence axiom by transforming objective probabilities into decision weights using a probability-weighting function.

The solution to this problem is to introduce a *second* sMPL that will allow estimation of the probability-weighting parameter representing the deviation from EU, in addition to the risk-aversion parameter.¹³ The problem of estimating the joint distribution of two such parameters with data from two sMPLs is the main subject of this paper.

2.2. Behavioural error models for sMPL data

When (conditional) independence can be assumed, there are two well-known candidate models for the analysis of choice under risk: the random preference (RP) model (Loomes and Sugden, 1998), which assumes that the preference parameter of interest varies randomly between tasks for a given subject; and the random utility (RU) model (Hey and Orme, 1994),¹⁴ which assumes that a subject uses the same preference parameter in every task, but makes errors in computing utilities (also known as “Fechnerian errors”) on each individual task. A third error model, known as the “tremble model” (Harless and Camerer, 1994), simply assumes that there is a small probability of a subject losing concentration and choosing randomly in any task, and can be used in conjunction with the previous two.

The peculiarity of the switching variant of MPL is that of imposing monotonicity of preferences. Interestingly, when the dependence structure (based on the assumption of a single switch-point) is fully incorporated into a model with within-subject variation, there is essentially only one decision per subject, and the model simplifies to the (univariate) MHP model that we have recommended. This was shown in Figure 1, panel (b), where the interval of values for r implied by the switch-point is strictly contained in all of the intervals implied by the choices, essentially implying that nearly all of the choice data is redundant. Essentially, using those choices does not add any information to knowledge of the interval of r between the two vertical lines.

It is worth comparing the plot in panel (b) of Figure 1 with that in panel (c), the latter showing the ranges of r implied by multiple switch-points. Both plots assume that the true value of r lies within the

¹²According to Andersen et al. (2006), “findings indicate that nothing is lost from using an enforced single switching point, as in the sMPL design” and “the [multiple] switching behavior that is often observed in MPL simply reflects indifference”.

¹³For the sake of completeness, we must mention the estimation approach used by HL in their seminal article, and by many others who followed in their footsteps. HL use the overall proportion of subjects choosing lottery A in each row across four MPLs, each representing one of four treatments. The parameters of a two-parameter Expo-Power Utility function (Saha, 1993) are estimated using maximum likelihood under the EU assumption. This method can be regarded as a Minimum-Distance estimator, fitting the data by selecting utility function parameters so that the estimated proportions of A choices across all rows and treatments align as closely as possible with the observed proportions. While this method provides a descriptive overview of the overall sample, it does not offer insights into individual behaviour or account for potential decision-making errors. Additionally, they report two sets of parameter estimates: one excluding multiple switchers and another including them. The differences in these estimates suggest that the subset of multiple switchers is not random, and excluding them would lead to biased estimates of the parameters of interest.

¹⁴A popular parameterisation of the Fechner model is the Luce (1959) formulation, applied by both Andersen et al. (2008) and Andreoni et al. (2015) in the joint estimation of risk and time preferences. These studies relied on the assumption of conditional independence between implied choices.

blue vertical bars, but while in the first case preferences are monotonic, in the second they are not. Such inconsistencies require the introduction of an error term. Any of the three above-mentioned models may be used, but for the RU and RP models the (conditional) independence assumption must hold. Both models would require the estimation of an additional parameter meant to accommodate the inconsistencies: the within-subject standard deviation of the preference parameter in the RP model; and that of the error made in computing utilities when the RU model is used. Whenever these two models are used with data from sMPL lists or from MPL lists which do not show multiple switching, there would not be sufficient information in the data to identify the additional error component.

When data shows multiple switch-points but there are reasons for believing that (conditional) independence does not hold, there is one other possible approach that could be used to estimate the preference parameter. It is based on the concept of “thick indifference curves” discussed by Bayrak and Hey (2020). In the interval-regression context, thick indifference curves amount to a widening of the interval known to contain the preference parameter. The notion of widening the interval to accommodate multiple switching is illustrated in panel (d) in Figure 1, where the interval of plausible values for r has been extended in order to make all the choices consistent with the observed switch-points. Larger intervals for r clearly result in more imprecise estimates, but accommodate the violation of independence between choices without introducing an additional error term to be estimated. This approach is particularly suitable when the proportion of subjects showing multiple switching is small. Its incidence in lab experiments is usually 5-10% (see, for example, Holt and Laury, 2002 and Brown and Healy, 2018). Another way of saying this is that, for the vast majority of subjects (at least 90%), the standard MPL and the sMPL are essentially the same.

3. Multiple sMPLs and the Multivariate Heterogeneous Preference model

Table 2 shows an example of a previously used two-sMPL design. This design was first used by Tanaka et al. (2010) and then by Riddell (2012), Liu (2013), Liu and Huang (2013), Riddell and Kolstoe (2013), Campos-Vazquez and Cuiilty (2014) and more recently by Bocquého et al. (2018). As with the Holt-Laury MPL, the lists display two alternative lotteries, but here there are 14 rows in each list. In both lists, all probabilities and outcomes are held fixed except the better outcome in lottery B. In both lists, the better outcome in lottery B increases through rows, making lottery B more and more appealing in terms of expected payoff. Note that the essential difference between the two lists is that in List 1 the probability of the high payoff is always low, while in List 2 the probability of the high payoff is always high. This difference is a valuable design feature in aiding the identification of the probability weighting parameters.

A subject presented with the two sMPLs shown in Table 2 will report one switch-point for each. In this section we explain how data on the two switch-points is used to estimate the Bivariate Heterogeneous Preference (BHP) model, that is, to estimate the parameters of the bivariate (between-subject) distribution of risk aversion and probability weighting.

Let the two lists in Table 2 be List 1 ($L = 1$) and List 2 ($L = 2$). Subject i , $i = 1, \dots, n$, faces the two lists, $L = 1, 2$. Each list comprises T rows indexed by t .¹⁵ The two lotteries appearing in row t are A^L and $B^{L,t}$. Let us denote the two outcomes of lottery A^L as a_1 and a_2 , with $a_1 > a_2$, occurring with probability p and $1 - p$, respectively. Similarly, the two outcomes of lottery $B^{L,t}$, are denoted as $b_1^{L,t}$ and b_2 , with $b_1^{L,t} > b_2 \forall t$, occurring with probability q and $1 - q$, respectively.

¹⁵The lists need not to contain the same number of rows. Here, for simplicity, we use a notation which applies strictly to the working case, but its generalisation is straightforward.

Table 2. The paired lottery choices used in Tanaka et al. (2010)

LIST 1									
Row	Lottery A				Lottery B				expected payoff difference
	probability (40)	payoff	probability (10)	payoff	probability (?)	payoff	probability (5)	payoff	
1	0.3	40	0.7	10	0.1	68	0.9	5	7.7
2	0.3	40	0.7	10	0.1	75	0.9	5	7.0
3	0.3	40	0.7	10	0.1	83	0.9	5	6.2
4	0.3	40	0.7	10	0.1	93	0.9	5	5.2
5	0.3	40	0.7	10	0.1	106	0.9	5	3.9
6	0.3	40	0.7	10	0.1	125	0.9	5	2.0
7	0.3	40	0.7	10	0.1	150	0.9	5	-5
8	0.3	40	0.7	10	0.1	185	0.9	5	-4.0
9	0.3	40	0.7	10	0.1	220	0.9	5	-7.5
10	0.3	40	0.7	10	0.1	300	0.9	5	-15.5
11	0.3	40	0.7	10	0.1	400	0.9	5	-25.5
12	0.3	40	0.7	10	0.1	600	0.9	5	-45.5
13	0.3	40	0.7	10	0.1	1,000	0.9	5	-85.5
14	0.3	40	0.7	10	0.1	1,700	0.9	5	-155.5
LIST 2									
Row	Lottery A				Lottery B				expected payoff difference
	probability (40)	payoff	probability (10)	payoff	probability (?)	payoff	probability (5)	payoff	
1	0.9	40	0.1	30	0.7	54	0.3	5	-0.3
2	0.9	40	0.1	30	0.7	56	0.3	5	-1.7
3	0.9	40	0.1	30	0.7	58	0.3	5	-3.1
4	0.9	40	0.1	30	0.7	60	0.3	5	-4.5
5	0.9	40	0.1	30	0.7	62	0.3	5	-5.9
6	0.9	40	0.1	30	0.7	65	0.3	5	-8.0
7	0.9	40	0.1	30	0.7	68	0.3	5	-10.1
8	0.9	40	0.1	30	0.7	72	0.3	5	-12.9
9	0.9	40	0.1	30	0.7	77	0.3	5	-16.4
10	0.9	40	0.1	30	0.7	83	0.3	5	-20.6
11	0.9	40	0.1	30	0.7	90	0.3	5	-25.5
12	0.9	40	0.1	30	0.7	100	0.3	5	-32.5
13	0.9	40	0.1	30	0.7	110	0.3	5	-39.5
14	0.9	40	0.1	30	0.7	130	0.3	5	-53.5

Note: In each list, subjects are asked to start choosing between lottery A (safer) and B (riskier) from the first row, and to proceed through the rows in order until lottery B is chosen. Prizes are expressed in 1,000 dong.

We assume that subjects are Rank Dependent Utility (RDU) maximisers. Accordingly, subject i evaluates the two lotteries, A^L and $B^{L,t}$, as

$$\begin{aligned}
 V_i(A^L) &= V_i(a_1, a_2, p; \alpha_i, \gamma_i) = w_i(p; \gamma_i) u_i(a_1; \alpha_i) + [1 - w_i(p; \gamma_i)] u_i(a_2; \alpha_i) \\
 V_i(B^{L,t}) &= V_i(b_1^{L,t}, b_2, q; \alpha_i, \gamma_i) = w_i(q; \gamma_i) u_i(b_1^{L,t}; \alpha_i) + [1 - w_i(q; \gamma_i)] u_i(b_2; \alpha_i) \quad (1)
 \end{aligned}$$

In (1), $u_i(z; \alpha_i)$ is the utility function, and for this we adopt the power functional form $u_i(z; \alpha_i) = z^{\alpha_i}$. The power parameter $\alpha_i > 0$ is less than 1 for risk-averse agents, equal to 1 for risk-neutral agents, and greater than 1 for risk-loving agents. $w_i(p; \gamma_i)$ is a probability-weighting function, where p is the true probability of the higher outcome. For this we adopt the one-parameter functional form proposed by Prelec (1998), $w_i(p; \gamma_i) = \exp[-(-\ln(p))^{\gamma_i}]$.¹⁶ The parameter $\gamma_i > 0$ determines the shape of the weighting function: $\gamma_i = 1$ implies no probability distortion, and the model reduces to EU; when $0 < \gamma_i < 1$, the probability weighting function takes on an inverse-s shape. In this case, subject i is attracted to low-probability high outcomes more than an EU subject would be. When $\gamma_i > 1$, the probability weighting function takes on an s-shape, so that subject i undervalues low-probability high outcomes.

Suppose that subject i decides to switch from lottery A to lottery B in row \tilde{t}^L of List $L \in \{1, 2\}$. Then, α_i and γ_i satisfy the following conditions, denoted as $C_i^L(\tilde{t}^L)$.

$$C_i^L(\tilde{t}^L) = \begin{cases} V_i(B^{L,1}) > V_i(A^L) & \Leftrightarrow \tilde{t}^L = 1 \\ V_i(B^{L,\tilde{t}^L}) > V_i(A^L) > V_i(B^{L,\tilde{t}^L-1}) & \Leftrightarrow 1 < \tilde{t}^L < T \\ V_i(B^{L,T}) < V_i(A^L) & \Leftrightarrow \tilde{t}^L = T \end{cases} \quad (2)$$

These conditions, applied to List 1 and List 2, produce the plots in Figure 2, top-left panel and top-right panel, respectively.¹⁷ The curved lines in the plots delimit the range of values for α_i and γ_i which jointly satisfy the conditions implied by i 's switch-point in the list. If i switches in row 1, then the implied pairs of values would be (for both lists) all those above the highest curve. If i switched in row 2, then the set of pairs of values of the parameters implied by that choice would be all those squeezed between the highest and second highest curves. This reasoning continues until we reach the lowest curve, which instead delimits from above all the values compatible with having never switched to lottery B. An important difference between the two Lists is that when the conditions (2) are applied to List 1, the delimiting curves are upward-sloping lines, whereas when applied to List 2, the curves are downward-sloping.

How insightful are these areas? Let us consider a subject who switches to lottery B in row 7 of List 1. Referring to the top-left panel of Figure 2, the pairs of values compatible with that choice are located between the 6th and 7th curve from the top. That area justifies the choices of an individual who is extremely risk averse and heavily overvalues unlikely high-gains (very small α and γ), one who is risk neutral and does not distort probabilities at all (both α and γ equal 1) and one who is extremely risk loving with an s-shaped probability weighting function (very high α and γ). These behaviours are all observationally equivalent in terms of List 1. A similar reasoning can be extended to all the

¹⁶ Obviously it would be possible to assume different weighting functions here. In particular, it might be considered more appropriate to assume a weighting function that is always above the 45-degree line, unlike the chosen Prelec (1998) function which exhibits a crossover point. The ‘‘optimist’’ weighting function extolled by Wilcox (2017); Wilcox (2023) is a strong candidate. Interestingly, when the lists presented in Table 2 are applied to the optimist weighting function, both lists produce upward-sloping curves in the grid, which amounts to a serious problem for the estimation of the joint distribution. A different pair of lists would be needed in order to produce a grid with the required mixture of upward- and downward-sloping curves. Since the main objective here is to demonstrate the estimation technique, we consider the choice over functional forms to be of secondary importance.

¹⁷ *Mutatis mutandis*, these same conditions apply to the list in Table 1 to compute the ranges of r . In that case, $w_i(p) = p$ such that the EU hypothesis substitutes that of RDU. Further details can be found in Appendix A.

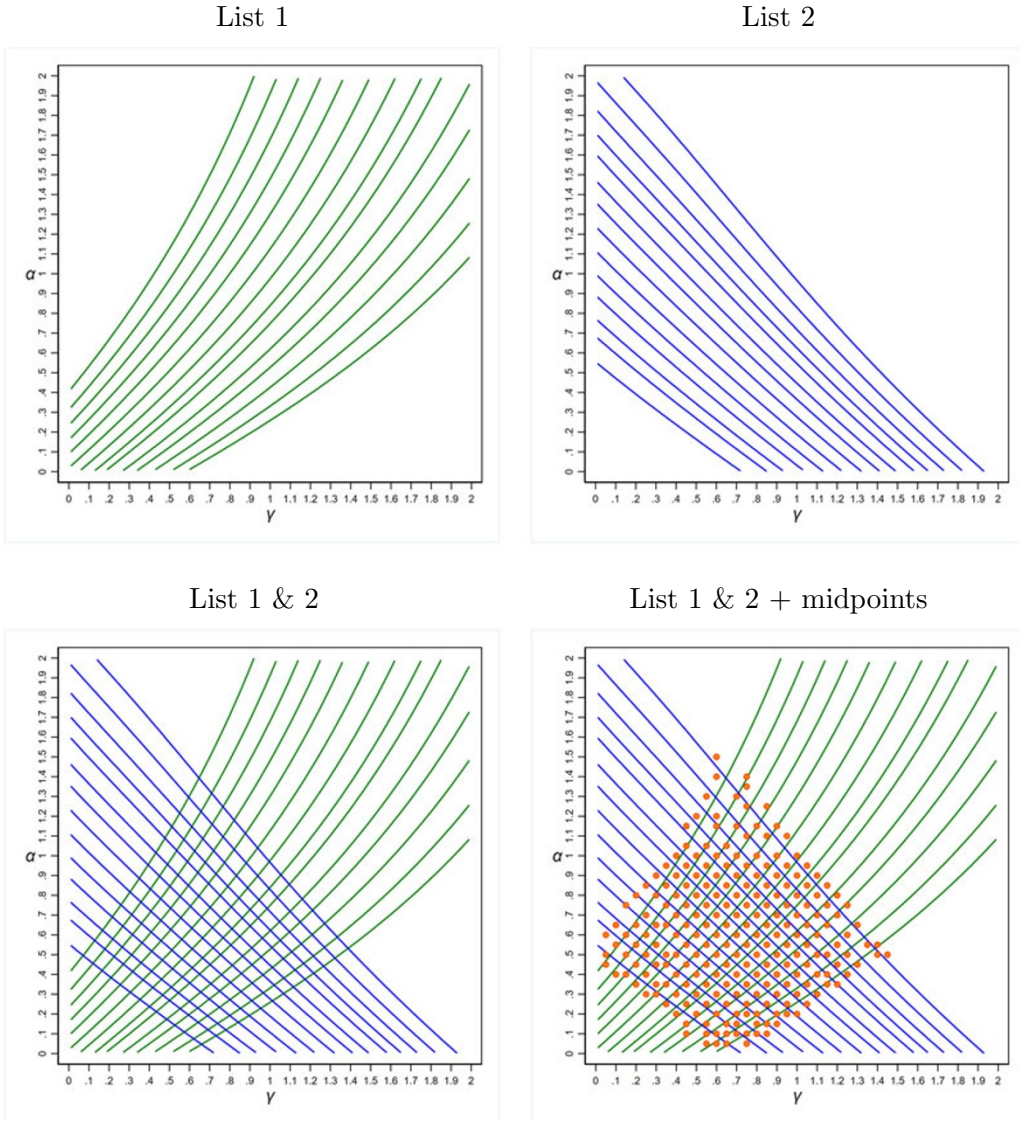


Figure 2. Sets of combinations of α and γ implied by switch-points for the sMPLs in Table 2
 Note: In all figures, the top and right bounds are infinite.

other switch-points of the two lists.¹⁸ Therefore, the two parameters are not identified from the sole information provided by a single list. This militates for the use of two lists together to circumscribe more tightly the area containing the values of the coefficients which apply to each type of behaviour under risk. On this point see also Drichoutis and Lusk (2016).

In effect, with the exception of the extreme choices (switching in the first row or not switching at all in one or both lists) conditions (2) for each pair of switch-points (one from List 1 and one from List 2) define a small, non-rectangular area where the parameters α and γ unequivocally lie, as shown in Figure 2, bottom-left panel.

¹⁸Crucially, switch-points in List 1 define areas where there exists a positive correlation between the values of the two parameters, while for List 2 such a correlation is negative.

Each of those areas contain information about subjects' attitude to risk and probability distortion. Developing an estimator that can exploit all of this information is the principal objective of this paper.

4. Estimation using midpoints

The estimation problem with data from two sMPLs has already been considered by Tanaka et al. (2010) and Liu (2013). In these studies, approximate values for the two parameters are obtained by taking the midpoint of the ranges defined for each parameter. These midpoints are then rounded to the nearest 0.05. The midpoints so calculated by Tanaka et al. (2010) are superimposed to the (α, γ) -areas in the bottom-right panel of Figure 2, and represented by the small orange circles. The cartesian coordinates of these midpoints (abscissae for γ and ordinates for α) are then used as the dependent variable in linear regressions, in order to investigate the impact of demographic or other variables on the preference parameters.¹⁹

While computationally straightforward, this approach has a number of shortcomings. For one, the method for obtaining the two parameter values can only be seen as a rough approximation, implying that the variables used in the analysis are certainly measured with error. Secondly, observations corresponding to a switch in row 1 or no-switch give rise to "open" and/or much larger intervals, for which the choice of midpoint is arbitrary. This can have an important effect on estimation, if a sizable proportion of the observations fall in these open intervals. Finally, and most importantly, the process of replacing intervals with midpoints results in the disregard of a significant amount of information that could be used to improve the precision of the parameter estimates.

On top of that, in the above-mentioned studies, the models for α and γ are estimated separately, and hence it is assumed that they are independently distributed. This assumption amounts to a further loss of information.²⁰ Figure 2 also makes clear that the estimation of one of the two coefficients cannot be separated from the estimation of the other. Any estimation strategy that incorporates all available information must be one in which the two preference parameters are modelled jointly. Therefore, the information provided by this graph concerning location and dimension of each area is crucial in order to make valid inference.

The approach detailed below addresses all these issues, since it estimates the joint distribution without introducing rounding and approximation error. In Appendix B, we also derive a modified version of the estimator for the joint distribution of α and γ based on midpoints. Through Monte Carlo simulation, we use this modified estimator as a countercheck of the validity of our estimator.

5. The Multivariate Heterogeneous Preference estimator with importance Sampling

As explained in Section 3, to be able to estimate the joint distribution of risk attitude and probability weighting, information is required from two distinct sMPLs for each subject. The principle underlying the estimator is that each pair of responses to the two-sMPL series corresponds to an irregularly -shaped region defined over the two parameters. The estimation procedure is based on finding the probability mass within this irregular (α, γ) -area.

By continuing with our working example, in what follows, we develop and evaluate our estimator in the bivariate case. It essentially consists of a generalisation of the interval regression model to two dimensions, but with non-rectangular areas. The estimator makes maximal use of the available information, and may therefore be classified as a full information estimator. As mentioned, we will refer to it as the Multivariate Heterogeneous Preference (MHP) estimator for sMPL data.

¹⁹Both studies also estimate by interval regression a loss-aversion coefficient from data collected via a third list, which needs not to be combined with the two lists we use in this study (see Table 2) to determine the ranges in which α and γ lie. For these two lists, both studies adopt the same utility and probability weighting functions assumed here.

²⁰Often, also the natural range of the parameters is overlooked by making a choice of the distribution to be fitted which extends to sets of "impossible" values, like, for example, when for a parameter only defined over the positive reals it is assumed a distribution with a domain which extends to the non-positive reals.

The two parameters of interest are described by a joint distribution with moments which may be allowed to vary by subject characteristics. Since α_i and γ_i are both required to be greater than 0, for the purpose of this exercise, we assume that they follow a bivariate lognormal distribution²¹

$$\begin{pmatrix} \alpha_i \\ \gamma_i \end{pmatrix} \sim \text{Lognormal} \left(\begin{pmatrix} \mu_\alpha \\ \mu_\gamma \end{pmatrix}, \begin{pmatrix} \sigma_\alpha^2 & \rho\sigma_\alpha\sigma_\gamma \\ \rho\sigma_\alpha\sigma_\gamma & \sigma_\gamma^2 \end{pmatrix} \right) \tag{3}$$

Here, μ indicates the mean, σ the standard deviation and ρ the correlation coefficient of the underlying bivariate normal distribution. The bivariate lognormal density function evaluated at (α_i, γ_i) is denoted as $f(\alpha_i, \gamma_i; \mu_\alpha, \sigma_\alpha, \mu_\gamma, \sigma_\gamma, \rho)$.

Given the distributional assumptions (3), we can define the probability of switching at $\tilde{t}^1 = h$ in List 1 and at $\tilde{t}^2 = k$ in List 2, with $h, k \in \{1, \dots, T\}$, as

$$\Pr(\tilde{t}^1 = h \cap \tilde{t}^2 = k) = \int_0^\infty \int_0^\infty \mathbf{1}[C_i^1(h) \cap C_i^2(k)] f(\alpha, \gamma) d\alpha d\gamma \tag{4}$$

where $\mathbf{1}[\cdot]$ is an indicator function taking the value 1 if the statement in brackets is true, 0 otherwise. The two conditions $C_i^1(h)$ and $C_i^2(k)$ appearing in (4) were defined in (2) above. Recall that these are the conditions which, when jointly satisfied, identify the (α, γ) -area corresponding to subject i 's switch-points (see Figure 2 above).

The two-dimensional integral (4) has no closed form solution, regardless of the distributional assumptions relating to α and γ . Moreover, the lines determined by each combination of switch-points are neither straight nor parallel, and define irregular areas having different sizes and shapes. Had those lines been both straight and parallel, we could have rotated the axes and estimated the parameters of interest using bivariate interval regression.²²

Analytical evaluation of the integral (4) is clearly infeasible and hence we have resorted to Monte Carlo integration. Choice of this method raises a number of further technical issues. First of all, let us consider the problem of sampling from any of the irregularly-shaped areas in Figure 2, bottom-left panel. Monte Carlo integration requires that, at each stage of the Maximum Simulated Likelihood procedure, at least one of the simulated draws is drawn from that particular area, otherwise it would result in a null likelihood contribution, preventing convergence. Using the Crude Frequency estimator,²³ which is the simplest and most widely used sampling procedure, this would not be guaranteed even by allowing a number of simulated draws per subject in the order of billions.²⁴ This is due to the so-called ‘‘rare event sampling’’ problem, which has been largely studied

²¹Alternatively, we could have used a bivariate normal truncated from below at 0 on both dimensions, for example, or any other bivariate distribution defined over the natural range of the parameters to be estimated.

²²Bivariate interval data has been encountered in areas other than Economics. Sun and Ding (2019) consider the timing of onset of bilateral eye disease. For each patient, onset of the disease in each eye is known to lie in an interval. To estimate the bivariate distribution of onset (i.e. joint distribution of onset in left and right eyes), a typical likelihood contribution is the volume of the bivariate distribution over a rectangle with horizontal and vertical sides. This feature of the likelihood function makes the estimation problem relatively straightforward, in contrast to the present application, for which the region over which integration is required is non-rectangular. See Moffatt (2021) for further detail relating to the Sun and Ding (2019) example. Although the bivariate interval regression model is not readily available in statistical packages, it would not be difficult to program for a number of popular bivariate distributions, given the rectangular areas over which integration is required. More on this in Appendix A.2.1.

²³This ‘‘unrefined’’ method consists in using a random number generator from the desired density. The procedure works well in a great number of cases but not always. For an approachable discussion of the most popular methods for drawing from densities, see Train (2009).

²⁴More details on this are available from the authors on request.

and discussed in the analysis of dynamical systems, in Physics, Computer Physics, Engineering and Biology applications, among others (see, for example, Rubinstein and Kroese 2017 and Morio et al. 2014).

Secondly, due to the presence of the indicator function in (4), the simulator is a step function in α and γ . This problem has been discussed extensively by McFadden (1984), McFadden and Ruud (1994), and Stern (1997).²⁵ Finally, the variance of the estimator is “unnecessarily large” (Stern, 1997).

Upon considering many other simulation techniques for drawing from a multivariate density (Acceptance-Rejection; Gibbs sampler; Metropolis-Hastings; and so on), we have settled on the choice of the Importance Sampling technique.²⁶ This is because of the appreciable similarity between the probability in (4) and that of the Multinomial Probit model (see Stern, 1997 [p. 2013, Eq. (2.21)]), which has enabled us to adapt to our case an estimator with very well-known properties, and also because with this technique the rare-event sampling problems can be easily addressed.

To understand the technique, it is helpful to rewrite the probability (4) of switching in row h of List 1 and in row k of List 2 as

$$\Pr(\tilde{t}^1 = h \cap \tilde{t}^2 = k) = \int_0^\infty \int_0^\infty \mathbf{1}[C_i^1(h) \cap C_i^2(k)] \frac{f(\alpha, \gamma)}{g(\alpha, \gamma)} g(\alpha, \gamma) d\alpha d\gamma \quad (5)$$

Here, we have innocuously divided and multiplied the integrand by $g(\alpha, \gamma)$, which is an easy-to-sample-from distribution in α and γ that is concentrated in an area where we know that the values of α and γ lie for subject i . This is the basic trick for drawing from densities referred to as “Importance Sampling”,²⁷ because it oversamples from the “important” part of the support of α and γ . Basically, this technique entails sampling from $g(\alpha, \gamma)$ instead of $f(\alpha, \gamma)$, and applying the “importance weight” $f(\alpha, \gamma)/g(\alpha, \gamma)$.

Let us see the Importance Sampling technique at work in our case. Suppose that subject i switches in row 4 of List 1 and row 5 of List 2. Figure 3 magnifies the irregularly-shaped area from Figure 2, bottom-left plot, that encloses all combinations of α and γ compatible with this choice. The red lines delineate a rectangle around that area which exactly contains it, establishing a lower and upper limit for the parameter α , that is $\alpha_{i,l}$ and $\alpha_{i,u}$, and for γ , that is $\gamma_{i,l}$ and $\gamma_{i,u}$, respectively. This rectangle is the “important” region of $\alpha - \gamma$ space for subject i . In order to circumvent the rare-event sampling issue discussed earlier, we choose $g(\alpha, \gamma)$ so that it has exactly this rectangle as support. Any bivariate distribution truncated to within the so-established limits can be used, and it is even possible to choose two different and independent distributions for the two dimensions. This ensures that we sample adequately from the curvilinear quadrilateral of interest.²⁸

Obviously, we have to address the fact that we are drawing from a limited area and not from the entire domain of α and γ . This is the purpose of the “importance weight” $f(\alpha, \gamma)/g(\alpha, \gamma)$ appearing in (5).

²⁵Stern (1997) writes: “Step functions are very difficult to optimize over because derivatives are equal to zero everywhere but points with discrete steps and do not exist at such points”.

²⁶We want to make clear that the above-mentioned estimation problems are not related to the choice of the distribution for the parameters (one can use common distributions like the normal and logistic, truncated in suitable ranges if necessary, or some less popular like the Johnson’s frequency curves (see Andersen et al., 2012), and so on), but result from the methods used to draw from such distributions, and in particular from the crude frequency estimator.

²⁷For a detailed description, see Gouriéroux and Monfort (1996).

²⁸This is not necessarily true for the open intervals. In those cases, the rectangles which frame the relevant (α, γ) -areas are only limited from below or from above. This is certainly a limitation, because the sampling area is not well restricted as for the closed (α, γ) -areas. However, this solution has proven to work particularly well also in those cases, because it restricts the sampling areas anyhow. In similar cases, in order to ensure that the pairs (α, γ) are sampled from the relevant open-ended areas, very intuitively, it is sufficient to make a choice for $g(\alpha, \gamma)$ that enables draws of positively correlated pairs of α and γ for upward-sloping open intervals and a different choice of $g(\alpha, \gamma)$ that enables draws of negatively correlated pairs of α and γ for downward-sloping open intervals.

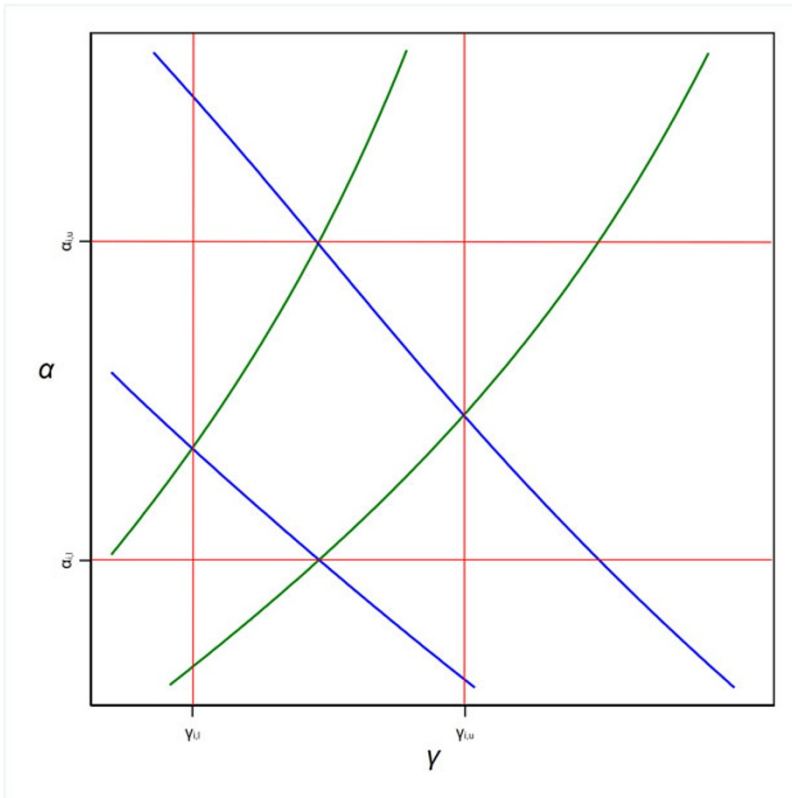


Figure 3. (α, γ) -range for a generic combination of switch-points and its bounds

5.1. The steps of Importance Sampling

Here, we give details on the steps of the procedure to simulate probability (5) by Importance Sampling:

- (1) Choose the number of draws, R . Draw $\hat{\alpha}_r$ and $\hat{\gamma}_r$, with $r = 1, \dots, R$, from the distribution $g(\alpha, \gamma)$ chosen according to the explained criteria; a good choice may be a distribution truncated within the rectangle that contains the (α, γ) -area we wish to sample;²⁹
- (2) compute the indicator $\mathbf{1}[C_i^1(h; \hat{\alpha}_r, \hat{\gamma}_r) \cap C_i^2(k; \hat{\alpha}_r, \hat{\gamma}_r)]$ following (1) and (2);
- (3) multiply by the importance weight $f(\hat{\alpha}_r, \hat{\gamma}_r) / g(\hat{\alpha}_r, \hat{\gamma}_r)$;
- (4) repeat all the previous steps R times and average the results.

An estimate of the probability of subject i switching in row h of List 1 and in row k of List 2 is given by:

$$\widehat{\Pr}(\tilde{t}^1 = h \cap \tilde{t}^2 = k) = \frac{\sum_{r=1}^R \mathbf{1}[C_i^1(h; \hat{\alpha}_r, \hat{\gamma}_r) \cap C_i^2(k; \hat{\alpha}_r, \hat{\gamma}_r)] \times f(\hat{\alpha}_r, \hat{\gamma}_r) / g(\hat{\alpha}_r, \hat{\gamma}_r)}{R} \quad (6)$$

As $R \rightarrow \infty$, (6) converges to the exact integral (5). (6) is the individual likelihood contribution that is passed to the likelihood function to maximise.³⁰ The so-obtained MHP estimator incorporates all

²⁹The distribution $g(\alpha, \gamma)$ need not be unique for all the (α, γ) -areas. For example, one may choose to draw from two independent truncated normals for the areas inside the thick grid and two correlated truncated normals for the surrounding open-ended areas.

³⁰The program is written in Stata.

the elements of the decision process via the conditions C_i^1 and C_i^2 which are evaluated at each step of the procedure.³¹

Importance Sampling estimators are proven to be consistent and to have a smaller variance than the crude frequency estimator. For an explanation and extensive discussion of the properties of this class of estimators see Stern (1997). Appendix D reports results from a Monte Carlo experiment which shows that the MHP estimator has good small-sample properties, while the midpoint estimator does not.

6. An application to real data

To demonstrate the application of our estimator to a real setting, we shall use the data set of Tanaka et al. (2010),³² whose design was presented in Table 2 above.

The results are shown in Table 3. For purposes of comparison, the results from estimation of both the MHP and the midpoint estimator are presented. For each of risk attitude and probability weighting, we report estimates of a model with no explanatory variables (specifications 1), and a model allowing both parameters to depend on some explanatory variables (specifications 2). Since our objective is simply to illustrate the application of our new technique, we only use two of the available explanatory variables, age and gender.

Both estimators provide evidence that the risk attitude parameter is affected negatively by age; that is, older subjects are more risk averse. Gender seems to have a mild negative effect on attitude to risk, only according to the midpoint estimator. Neither age nor gender seem to affect probability weighting. However, in all cases the estimates of the standard deviations of $\log(\alpha)$ and $\log(\gamma)$ differ greatly between the two methods. Note also that the estimate of the correlation coefficient is approximately twice as large when MHP is used.

It is worth noting that, even if the estimates of μ_α and μ_γ do not differ dramatically between the two methods, in all cases the midpoint estimator produces estimated standard errors which are notably smaller than those associated with the estimates from the MHP estimator. This results in the higher apparent significance of the midpoint parameter estimates (e.g. the effect of gender on risk attitude). This finding is consistent with the higher actual sizes obtained for the midpoint estimator in the Monte Carlo results reported in Appendix D. That finding essentially tells us when the midpoint estimator is used there is a tendency for significant effects to be detected when no effect is present.

We might be interested in estimating the mean risk attitude and the mean probability distortion from the results in Table 3. We use specifications (1) to keep it simple. To do this for our parameters which are assumed to follow a lognormal distribution, we have computed the following formulae for α and γ , respectively, as explained in footnote 5 in the Appendix,

$$\exp\left(\hat{\mu}_\alpha + \frac{\hat{\sigma}_\alpha^2}{2}\right) \quad \exp\left(\hat{\mu}_\gamma + \frac{\hat{\sigma}_\gamma^2}{2}\right) \quad (7)$$

where the hat indicates parameter estimates.

The results of this exercise are reported in Table 4. Interestingly, the MHP estimates of the means of α and γ are both notably larger than the midpoint estimates. In effect, while the midpoint estimation gives rise to what is normally observed from data, that is risk aversion and probability distortion, the MHP estimation produces estimates that are much closer to 1 (the mean of α is not significantly different from 1; that of γ is still somewhat below 1). However, the null hypothesis that the means of α and γ are jointly equal to 1 (which would imply that the average subject is an expected-value maximiser) is rejected (Wald test: $\chi^2(2) = 11.77$, p -value = 0.0028).

³¹We note that the lines that delimit each curvilinear quadrilateral (see the example in Figure 3) are not explicitly derived and/or incorporated in the estimator. Those lines are implicitly drawn by the conditions in (2) which incorporate the lottery evaluations in (1).

³²Available at <https://www.openicpsr.org/openicpsr/project/112336/version/V1/view>.

Table 3. Estimation results of the two models from Tanaka et al.'s data

	MHP		Midpoint estimator	
	(1)	(2)	(1)	(2)
μ_α				
Age		-0.0147** (0.0065)		-0.0103** (0.0045)
Gender		-0.2368 (0.1745)		-0.2285* (0.1196)
Constant	-0.6941*** (0.0868)	-0.1157 (0.2351)	-0.7272*** (0.0595)	-0.2832** (0.1475)
σ_α	1.1360*** (0.0690)	1.1140*** (0.0673)	0.8005*** (0.0421)	0.7813*** (0.0411)
μ_γ				
Age		-0.0044 (0.0033)		-0.0026 (0.0024)
Gender		-0.1032 (0.0881)		-0.0982 (0.0648)
Constant	-0.3070*** (0.0432)	-0.1148 (0.1186)	-0.3805*** (0.0318)	-0.2418*** (0.0799)
σ_γ	0.5688*** (0.0336)	0.5635*** (0.0333)	0.4274*** (0.0225)	0.4233*** (0.0222)
ρ	0.4197*** (0.0698)	0.4049*** (0.0710)	0.2272*** (0.0704)	0.2049*** (0.0712)
Loglikelihood	-1017.4258	-1013.6761	-314.7112	-309.5042
observations	181	181	181	181

Note: Standard errors in parentheses. * p -value < .10, ** p -value < .05, *** p -value < .01. We note that the values of the log-likelihoods of the two models cannot be compared because both the dependent variables and the structure of the likelihood functions are different.

The reason for such a difference is likely to be found in the transformation needed to get the means of α and γ (see (7)), which exponentiates the sum of μ and $\sigma^2/2$. As a consequence, if both μ and σ are negatively biased, the mean of α and/or γ may result heavily underestimated; the opposite would hold if μ and σ were both positively biased. What would happen if one parameter is estimated with a positive bias and the other with a negative bias depends on the relative dimension of such distortions.

Obviously, we cannot say which of the two sets of estimates approaches more closely the true values of the parameters, since we do not know the true DGP. However, we can make some conjectures based on the simulation results presented in Appendix D. The peculiarity of this data set lies in the distribution of the data points over the grid: 45% of the observations are positioned outside the thick grid (for the experimental subjects having switched at the first row or never switched in one or both lists) and the remainder are heavily clustered at the very centre of the grid (corresponding to switch-points 5, 6, 7 and 8). The characteristics of the sample, the estimates of the means of α and γ and the respective underlying parameters resemble the DGP4 case in Table D.1 discussed in the Appendix. That is a limiting case in which the midpoint estimator pronouncedly underestimates all the parameters with respect to the MHP estimator, but inference provided from both estimators proves unreliable.

These findings confirm the importance of using a suitable experimental design along with a valid estimator.

Table 4. Estimates of the means of α and γ from estimation results in Table 3

Estimated mean of α		Estimated mean of γ	
(MHP)	(midpoint)	(MHP)	(midpoint)
0.9523	0.6658	0.8649	0.7489
[0.7420; 1.1625]	[0.5765; 0.7550]	[0.7856; 0.9442]	[0.7001; 0.7976]

Note: 95% confidence intervals obtained by the delta method in parentheses.

7. Extensions, recommendations to practitioners, and words of warning

One obvious attraction of our MHP estimator is that it consistently estimates the parameters of the joint distribution of the preference parameters, thereby capturing unexplained heterogeneity in these parameters. Another attraction is that it can also (as demonstrated in Table 3) allow for explained heterogeneity, by allowing the means of the preference parameters to depend linearly on explanatory variables (e.g. treatments, demographics).

There are other reasons why knowledge of these distributional parameters is very useful. For example, knowledge of these parameters is essential if formal methods of optimal design are to be used in developing a fully efficient design for a future experiment.

Yet another attraction of our approach is that, in post-estimation, a set of subject-specific posterior estimates of preference parameters may be computed by applying Bayes' Rule. Let us consider how such posterior estimates may be obtained. Equation (6) above presented the likelihood contribution for a subject switching in rows h and k . Having estimated the model, we can obtain posterior estimates of the two parameters, again for a subject switching in rows h and k , using:

$$\hat{E}(\alpha_i | \tilde{t}_i^1 = h \cap \tilde{t}_i^2 = k) = \frac{\sum_{r=1}^R \hat{\alpha}_r \mathbf{1}[C_i^1(h; \hat{\alpha}_r, \hat{\gamma}_r) \cap C_i^2(k; \hat{\alpha}_r, \hat{\gamma}_r)] \times f(\hat{\alpha}_r, \hat{\gamma}_r) / g(\hat{\alpha}_r, \hat{\gamma}_r)}{\sum_{r=1}^R \mathbf{1}[C_i^1(h; \hat{\alpha}_r, \hat{\gamma}_r) \cap C_i^2(k; \hat{\alpha}_r, \hat{\gamma}_r)] \times f(\hat{\alpha}_r, \hat{\gamma}_r) / g(\hat{\alpha}_r, \hat{\gamma}_r)} \quad (8)$$

and:

$$\hat{E}(\gamma_i | \tilde{t}_i^1 = h \cap \tilde{t}_i^2 = k) = \frac{\sum_{r=1}^R \hat{\gamma}_r \mathbf{1}[C_i^1(h; \hat{\alpha}_r, \hat{\gamma}_r) \cap C_i^2(k; \hat{\alpha}_r, \hat{\gamma}_r)] \times f(\hat{\alpha}_r, \hat{\gamma}_r) / g(\hat{\alpha}_r, \hat{\gamma}_r)}{\sum_{r=1}^R \mathbf{1}[C_i^1(h; \hat{\alpha}_r, \hat{\gamma}_r) \cap C_i^2(k; \hat{\alpha}_r, \hat{\gamma}_r)] \times f(\hat{\alpha}_r, \hat{\gamma}_r) / g(\hat{\alpha}_r, \hat{\gamma}_r)} \quad (9)$$

Posterior estimates obtained using (8) and (9) can then be used as independent variables in the behavioural model of central interest. Gillen et al. (2019) has raised the issue of measurement error in this context. Note that estimates obtained using (8) and (9) are superior to, and less susceptible to measurement error than, midpoint estimates that have been used in previous research. This is a particular advantage in the case of “open intervals” because the midpoint of an open-interval must be assigned arbitrarily.

Although the current specification of the MHP model assumes only continuous variation in the preference parameters, it would be possible to extend the model to allow for a positive probability mass for different types, e.g. EU- and EV-maximisers. The extended model would have the flavour of a finite mixture model, made popular in the risky choice context by Harrison and Rutström (2009) and Conte et al. (2011). This seems a promising avenue for further research.

It would also be possible to estimate the MHP model using a different specification for utility, e.g. CARA rather than CRRA, and/or a different specification for probability weighting function, e.g. power, Tversky-Kahneman, etc. It is of course possible that the experimental design of the MPLs would need to change to accommodate such changes in specification (see footnote 16 previously in this article). Competing specifications could be compared using suitable non-nested testing procedures.

Regarding recommendations, we should point out that one convincing reason for using an sMPL is that it avoids within-subject variation, thereby avoiding the need to account for it. Of course, if there

are more sMPLs than parameters, we allow within-subject variation to return. However, we must say that avoiding it by imposing a single switch-point and then reintroducing it by using more lists seems illogical. If a researcher is in the position of finding that they have more lists than parameters of interest, then one way of allowing for inconsistencies is to apply the concept of “thick indifference curves” discussed in Section 2.2.

Moreover, we strongly advise researchers to choose between two approaches at the start of their analysis:

- (1) If their focus is on behaviour under risk, or closely associated research areas such as testing the axioms of rationality, testing non-EU theories, estimating within-subject variation, investigating the effect of decision-making experience on risk-taking, and so on, then they should design their experiment as a sequence of choice problems, presented independently. The set of choice problems do not need to be part of any list. The approach to choosing the problems might be based on the requirement to cover the whole of the Marschak-Machina Triangle, as is the approach of Hey and Orme (1994), Loomes and Sugden (1998), Hey (2001) and others. The choice of design could be improved further by optimising the design in the statistical sense.³³ Having applied such an experimental design, the resulting data set should be analysed using a fully structural choice model, assuming both between- and within-subject variation in behaviour.
- (2) If their focus is on some other aspect of economic behaviour, in which risk attitude and other preference parameters might play an important explanatory role, the elicitation of these preference parameters must surely take only a small part of the time allocated to the experiment. For this purpose, one or two (or possibly three) sMPLs should be used, depending on the number of preference parameters of interest. Note that the sMPL should be chosen in preference to the MPL, to avoid completely the problem of multiple switching. Having obtained the data on switch-points, we would obviously recommend the use of our MHP model to estimate the joint distribution of preference parameters, and to deduce (via Bayes’ rule, see (8) and (9) above) a high-quality estimate of each preference parameter for each individual subject. These estimates could then be used as valid explanatory variables in the model of central interest.

We emphasise that we are advising researchers to choose one of these two extreme approaches, depending on their research objectives. In each of the two approaches, there is a clearly defined and internally consistent econometric modelling strategy to follow. We strongly advise avoiding any approach that lies between the two extremes, since the resulting data set will not, for reasons given throughout this paper, comply with any valid econometric modelling strategy: it will require either incorrect assumptions such as the independence between implied choices, or *ad hoc* assumptions such as that of “thick” indifference curves.

Researchers choosing to go down path (2) above should be reminded of an important caveat. The sMPL approach is based on the key assumption of deterministic revealed preference, since this is what enables us to assume away within-subject variation in preferences.³⁴ We must acknowledge that there is an extensive experimental literature calling this assumption into question (usually based on evidence of imperfect retest reliability of repeated trials of identical decision or pricing problems).³⁵

If a practitioner is determined to allow for a violation of the assumption of deterministic revealed preference when using sMPLs, we would advise them to treat the posterior estimates extracted from

³³See Moffatt (2015), Chapter 14, who considers the problem of finding a D-optimal design in the context of a risky choice experiment.

³⁴One of the Reviewers describes the assumption of deterministic revealed preference elegantly as “the deadest of dead parrots”.

³⁵For example Ballinger and Wilcox (1997), Camerer (1989), Hey and Orme (1994), Butler and Loomes (2007), Loomes and Pogrebná (2014), Loomes and Sugden (1998).

the MHP model (using (8) and (9) above) as measures of true preferences that are subject to classic measurement error. They could then follow Gillen et al. (2019) by using instrumental variables approaches to correct for the bias caused by measurement error. All that would be required is some other preference measure that could be used as an instrument.

8. Conclusion

The MPL choice series and its “switching” variant are popular tools for eliciting preferences (often risk-attitude, but also time preferences, ambiguity aversion, loss aversion, and so on) from subjects (see footnotes 1, 2 and 3). The framework is considered particularly useful in the following settings: i) when preferences for risk are elicited as a part of an experiment that is not strictly focused on behaviour under risk but requires some control for risk attitude; ii) in the field, where the experimental subject is not acquainted with and ready for long experimental procedures. In fact, MPL and sMPL have been chosen for experiments run in developing countries by Tanaka et al. (2010), among others, where facing often illiterate subjects with a long list of random lottery pairs would have been infeasible.

We must also acknowledge, though, that some researchers (see, among others, Bosch-Domènech and Silvestre, 2013, and Freeman et al., 2019) maintain that these tools may have some limitations, especially when the aim of the study is that of testing particular axioms relating to the preference parameter being measured. And, to repeat the point made at the end of the last section, the framework relies on the assumption of deterministic revealed preference which must be considered questionable.

Despite these limitations, it is an unavoidable fact that the MPL/sMPL approach is widely used, and highly popular with practitioners. Moreover, given the widely documented inadequacy of EU to explain empirical observations, it is becoming common for designs to include more than one sMPL, enabling simultaneous estimation of more than one preference parameter in line with non-EU theories. Here is where our contribution lies, since we are developing a valid estimation procedure in settings with more than one sMPL.

We have named our estimator the Multivariate Heterogeneous Preference (MHP) estimator. Implementation of the approach has proved challenging econometrically, both at the design stage, and the estimation stage. At the design stage, we face the problem that the optimal design of the sMPLs depends on the number of preference parameters of interest, and on the choice of specifications in which the preference parameters of interest appear (see footnote 16). At the estimation stage, we face significant challenges in evaluating the likelihood function, and we have overcome these challenges using Monte Carlo integration with importance sampling.

We have applied the estimation approach to the bivariate case of Rank Dependent Utility for choices under risk. However, the realms of applicability are much wider. For example, the approach could be applied to multi-parameter models of discount rates, or to social preference models such as the Fehr and Schmidt model (see Fehr and Schmidt, 1999).

The extension to more than two sMPLs is clearly an alluring prospect because it allows the estimation of the joint distribution of more than two parameters, or alternatively allows the incorporation of within-subject variation. Moreover, the MHP estimator can be adapted to estimate preferences from choice tasks which have been presented to subjects one-at-a-time in a random order.

Using Monte Carlo simulation, we have investigated the statistical properties of the MHP estimator, and we have established that it is eminently superior to the midpoint estimator and its variants, which have made several appearances in the literature.

Finally, we have applied our MHP estimator to a real data set and compared the results to those obtained from the midpoint estimator. We have confirmed that the two estimators are not interchangeable, and may lead to very different conclusions.

Far be it from us to maintain that this is the “best” possible estimator of this sort of data, we can certainly say that our estimator is useful, outperforms the popular midpoint estimator, and, we hope,

may initiate the development of further competing multivariate estimators for the most effective use of MPL and sMPL data, whose adequacy to measure individual preferences cannot be properly judged and should not be assessed if not via good-property estimators.

References

- Abdellaoui, M., Driouchi, A., & L'Haridon, O. (2011). Risk aversion elicitation: Reconciling tractability and bias minimization. *Theory and Decision*, 71(1), 63–80.
- Alonso, J., Di Paolo, R., Ponti, G., & Sartarelli, M. (2018). Facts and misconceptions about 2D:4D, social and risk preferences. *Frontiers in Behavioral Neuroscience*, 12, 22.
- Andersen, S., Harrison, G., Hole, A., Lau, M., & Rutström, E. (2012). Non-linear mixed logit. *Theory and Decision*, 73(1), 77–96.
- Andersen, S., Harrison, G., Lau, M., & Rutström, E. (2006). Elicitation using multiple price list formats. *Experimental Economics*, 9(4), 383–405.
- Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2008). Eliciting risk and time preferences. *Econometrica*, 76(3), 583–618.
- Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2014). Discounting behavior: A reconsideration. *European Economic Review*, 71, 15–33.
- Andersson, O., Holm, H. J., Tyran, J. -R., & Wengström, E. (2020). Robust inference in risk elicitation tasks. *Journal of Risk and Uncertainty*, 61(3), 195–209.
- Andreoni, J., Kuhn, M. A., & Sprenger, C. (2015). Measuring time preferences: A comparison of experimental methods. *Journal of Economic Behavior & Organization*, 116, 451–464.
- Attanasi, G., Georgantzis, N., Rotondi, V. & Vigani, D. (2018). Lottery- and survey-based risk attitudes linked through a multichoice elicitation task. *Theory and Decision*, 84(3), 341–372.
- Aycinena, D., Baltaduonis, R. & Rentschler, L. (2014). Risk preferences and prenatal exposure to sex hormones for ladinos. *PLOS ONE*, 9(8), 1–10.
- Åstebro T., Mata, J. & Santos-Pinto, L. (2014). Skewness seeking: Risk loving, optimism or overweighting of small probabilities?. *Theory and Decision*, 78(2), 189–208.
- Ballinger, T. P. & Wilcox, N. T. (1997). Decisions, error and heterogeneity. *Economic Journal*, 107(443), 1090–1105.
- Bardsley, N., Cubitt, R., Loomes, G., Moffatt, P., Starmer, C. and Sugden, R. (2010). *Experimental economics: rethinking the rules*. Princeton University Press.
- Bartczak, A., Chilton, S., Czajkowski, M. & Meyerhoff, J. (2017). Gain and loss of money in a choice experiment. the impact of financial loss aversion and risk preferences on willingness to pay to avoid renewable energy externalities. *Energy Economics*, 65, 326–334.
- Bayrak, O. K. & Hey, J. D. (2020). Understanding preference imprecision. *Journal of Economic Surveys*, 34(1), 154–174.
- Beauchamp, J. P., Benjamin, D. J., Laibson, D. I., & Chabris, C. F. (2020). Measuring and controlling for the compromise effect when estimating risk preference parameters. *Experimental Economics*, 23(444), 1069–1099.
- Bellemare, C. & Shearer, B. (2010). Sorting, incentives and risk preferences: Evidence from a field experiment. *Economics Letters*, 108(3), 345–348.
- Binswanger, H. P. (1980). Attitudes toward risk: Experimental measurement in rural India. *American Journal of Agricultural Economics*, 62(3), 395–407.
- Bocquého, G., Deschamps, M., Helstroffer, J. and Joxhe, M. (2018) *Risk and refugee migration*. CREA Discussion Paper Series 18-08, Center for Research in Economic Analysis University of Luxembourg.
- Bocquého, G., Jacquet, F. & Reynaud, A. (2013). Expected utility or prospect theory maximisers? Assessing farmers' risk behaviour from field-experiment data. *European Review of Agricultural Economics*, 41(1), 135–172.
- Bolton, G. E., Ockenfels, A. & Stauf, J. (2015). Social responsibility promotes conservative risk behavior. *European Economic Review*, 74, 109–127.
- Bonan, J., LeMay-Boucher, P. & Scott, D. (2022). Can hypothetical measures of time preference predict actual and incentivised behaviour? Evidence from senegal. *World Development*, 159, 106029.
- Bonjean, I. (2023). Who are the loss-averse farmers? Experimental evidence from structurally estimated risk preferences. *European Review of Agricultural Economics*, 50(2), 421–456.
- Bosch-Domènech, A. & Silvestre, J. (2013). Measuring risk aversion with lists: A new bias. *Theory and Decision*, 75(4), 465–496.
- Bougherara, D., Friesen, L. & Nauges, C. (2021). Risk taking with left- and right-skewed lotteries. *Journal of Risk and Uncertainty*. 62(1), 81–112.
- Bougherara, D., Gassmann, X., Piet, L. & Reynaud, A. (2017). Structural estimation of farmers' risk and ambiguity preferences: A field experiment. *European Review of Agricultural Economics*, 44(5), 782–808.
- Boyd, C. M. & Bellemare, M. F. (2022). Why not insure prices? Experimental evidence from peru. *Journal of Economic Behavior & Organization*, 202, 580–631.

- Bradford, W. D., Dolan, P. & Galizzi, M. M. (2019). Looking ahead: Subjective time perception and individual discounting. *Journal of Risk and Uncertainty*, 58(1), 43–69.
- Brañas-Garza, P., Jorrat, D., Espín, A. M. and Sánchez, A. (2022). Paid and hypothetical time preferences are the same: Lab, field and online evidence. *Experimental Economics*, 26(2), 412–434.
- Brañas-Garza, P. & Rustichini, A. (2011). Organizing effects of testosterone and economic behavior: Not just risk taking. *PLOS ONE*, 6(12), 1–8.
- Brown, A. L. & Healy, P. J. (2018). Separated decisions. *European Economic Review*, 101, 20–34.
- Bruhin, A., Fehr-Duda, H. & Epper, T. (2010). Risk and rationality: Uncovering heterogeneity in probability distortion. *Econometrica*, 78(4), 1375–1412.
- Bruner, D. M. (2017). Does decision error decrease with risk aversion?. *Experimental Economics*, 20(1), 259–273.
- Butler, D. J. & Loomes, G. C. (2007). Imprecision as an account of the preference reversal phenomenon. *American Economic Review*, 97(1), 277–297.
- Camerer, C. F. (1989). An experimental test of several generalized utility theories. *Journal of Risk and Uncertainty*, 2(1), 61–104.
- Campos-Vazquez, R. & Cuijty, E. (2014). The role of emotions on risk aversion: A Prospect Theory experiment. *Journal of Behavioral and Experimental Economics*, 50, 1–9.
- Charness, G., Eckel, C. C., Gneezy, U. & Kajackaite, A. (2018). Complexity in risk elicitation may affect the conclusions: A demonstration using gender differences. *Journal of Risk and Uncertainty*, 56(1), 1–17.
- Charness, G., Garcia, T., Offerman, T. & Villeval, M. -C. (2020). Do measures of risk attitude in the laboratory predict behavior under risk in and outside of the laboratory?. *Journal of Risk and Uncertainty*, 60(2), 99–123.
- Cheung, S. L. (2020). Eliciting utility curvature in time preference. *Experimental Economics*, 23(2), 493–525.
- Coller, M. & Williams, M. (1999). Eliciting individual discount rates. *Experimental Economics*, 2(2), 107–127.
- Conte, A., Hey, J. & Moffatt, P. (2011). Mixture models of choice under risk. *Journal of Econometrics*, 162(1), 79–88.
- Crosetto, P. & Filippin, A. (2016). A theoretical and experimental appraisal of four risk elicitation methods. *Experimental Economics*, 19(3), 613–641.
- Csermely, T. & Rabas, A. (2016). How to reveal people's preferences: Comparing time consistency and predictive power of multiple price list risk elicitation methods. *Journal of Risk and Uncertainty*, 53(2), 107–136.
- Danz, D., Vesterlund, L. & Wilson, A. J. (2022). Belief elicitation and behavioral incentive compatibility. *American Economic Review*, 112(9), 2851–2883.
- Danz, D., Vesterlund, L. & Wilson, A. J. (2024). Evaluating behavioral incentive compatibility: Insights from experiments. *Journal of Economic Perspectives*, 38(4), 131–154.
- Drichoutis, A. C., & Lusk, J. (2016). What can multiple price lists really tell us about risk preferences?. *Journal of Risk and Uncertainty*, 53, 89–106.
- Drichoutis, A. C. & Nayga, R. M. J. (2015). Do risk and time preferences have biological roots? *Southern Economic Journal*, 82(1), 235–256.
- Ebert, S. & Wiesen, D. (2014). Joint measurement of risk aversion, prudence, and temperance. *Journal of Risk and Uncertainty*, 48(3), 231–252.
- Eckel, C. C. & 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., Johnson, C., & Montmarquette, C. (2004). *Saving decisions of the working poor: Short- and long-term horizons. Research in Experimental Economics*, 10, 219–260.
- Eijkelenboom, G. G., Rohde, I., & Vostroknutov, A. (2019). The impact of the level of responsibility on choices under risk: the role of blame. *Experimental Economics*, 22(4), 794–814.
- Fehr, E. & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3), 817–868.
- Ferecatu, A. & Öncüler, A. (2016). Heterogeneous risk and time preferences. *Journal of Risk and Uncertainty*, 53(1), 1–28.
- Freeman, D. J., Halevy, Y. & Kneeland, T. (2019). Eliciting risk preferences using choice lists. *Quantitative Economics*, 10(1), 217–237.
- Garbarino, E., Slonim, R. & Sydnor, J. (2010). Digit ratios (2D:4D) as predictors of risky decision making for both sexes. *Journal of Risk and Uncertainty*, 42(1), 1–26.
- Gillen, B., Snowberg, E. & Yariv, L. (2019). Experimenting with measurement error: Techniques with applications to the caltech cohort study. *Journal of Political Economy*, 127(4), 1826–1863.
- Goldzahl, L. (2017). Contributions of risk preference, time orientation and perceptions to breast cancer screening regularity. *Social Science & Medicine*, 185, 147–157.
- Gonzalez, R. & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology*, 38(1), 129–166.
- Gouriéroux, C. and Monfort, A. (1996). *Simulation-based Econometric Methods. OUP/CORE Lecture Series*. Oxford University Press.
- Habib, S., Friedman, D., Crockett, S. & James, D. (2017). Payoff and presentation modulation of elicited risk preferences in MPLs. *Journal of the Economic Science Association*, 3(2), 183–194.

- Hammersley, J. M. and Handscomb, D. C. (1964). *Monte carlo methods*. Methuen & Co.
- Harless, D. W. & Camerer, C. F. (1994). The predictive utility of Generalized Expected Utility theories. *Econometrica*, 62(6), 1251–1289.
- Harrison, G. W., & Rutström, E. (2008). Risk aversion in the laboratory. *Research in Experimental Economics*, 12, 41–196.
- Harrison, G. W. & Rutström, E. E. (2009). Expected utility theory and prospect theory: One wedding and a decent funeral. *Experimental economics*, 12(2), 133–158.
- He, H. & Villeval, M. C. (2017). Are group members less inequality averse than individual decision makers?. *Journal of Economic Behavior & Organization*, 138, 111–124.
- He, T. -S. & Hong, F. (2018). Risk breeds risk aversion. *Experimental Economics*, 21(4), 815–835.
- Hey, J. (2001). Does repetition improve consistency? *Experimental Economics*, 4(1), 5–54.
- Hey, J. D. & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica*, 62(6), 1291–1326.
- Holt, C. A. & Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5), 1644–1655.
- Isoni, A., Brooks, P., Loomes, G. & Sugden, R. (2016). Do markets reveal preferences or shape them?. *Journal of Economic Behavior & Organization*, 122, 1–16.
- Jack, B. K., McDermott, K. & Sautmann, A. (2022). Multiple price lists for willingness to pay elicitation. *Journal of Development Economics*, 159, 102977.
- Kaas, K. P. & Ruprecht, H. (2006). Are the Vickrey Auction and the BDM Mechanism Really Incentive Compatible? - Empirical Results and Optimal Bidding Strategies in Cases of Uncertain Willingness-to-pay. *Schmalenbach Business Review (sbr)*, 58(1), 37–55.
- Kahneman, D., Knetsch, J. L. & Thaler, R. H. (1990). Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy*, 98(6), 1325–1348.
- Kreft, C., Huber, R., Wuepper, D. & Finger, R. (2021). The role of non-cognitive skills in farmers' adoption of climate change mitigation measures. *Ecological Economics*, 189, 107169.
- Liebenehm, S. & Waibel, H. (2014). Simultaneous estimation of risk and time preferences among small-scale cattle farmers in west africa. *American Journal of Agricultural Economics*, 96(5), 1420–1438.
- Liu, E. M. (2013). Time to change what to sow: Risk preferences and technology adoption decisions of cotton farmers in China. *The Review of Economics and Statistics*, 95(4), 1386–1403.
- Liu, E. M. & Huang, J. (2013). Risk preferences and pesticide use by cotton farmers in China. *Journal of Development Economics*, 103, 202–215.
- Loomes, G. & Pogrebna, G. (2014). Measuring individual risk attitudes when preferences are imprecise. *The Economic Journal*, 124, 569–593.
- Loomes, G., Starmer, C. & Sugden, R. (2003). Do anomalies disappear in repeated markets?. *The Economic Journal*, 113(486), C153–C166.
- Loomes, G., Starmer, C. & Sugden, R. (2010). Preference reversals and disparities between willingness to pay and willingness to accept in repeated markets. *Journal of Economic Psychology*, 31(3), 374–387.
- Loomes, G. & Sugden, R. (1998). Testing different stochastic specifications of risky choice. *Economica*, 65(260), 581–598.
- Luce, R. D. (1959). *Individual Choice behavior: A Theoretical analysis*. John Wiley and Sons.
- Martinez, S. -K., Meier, S. & Sprenger, C. (2023). Procrastination in the Field: Evidence from Tax Filing. *Journal of the European Economic Association* 21(3), 1119–1153.
- McFadden, D. L., & Ruud, P. A. (1994). Estimation by simulation. *The Review of Economics and Statistics*, 76(4), 591–608.
- McFadden, D. L. (1984). Econometric analysis of qualitative response models. In Z. Griliches & M. D. Intriligator (Eds.), *Handbook of econometrics*. (vol 2, chapter 24, pp. 1395–1457), Elsevier.
- Meissner, T., Gassmann, X., Faure, C. & Schleich, J. (2023). Individual characteristics associated with risk and time preferences: A multi country representative survey. *Journal of Risk and Uncertainty*, 66(1), 77–107.
- Menapace, L., Colson, G. & Raffaelli, R. (2013). Risk aversion, subjective beliefs, and farmer risk management strategies. *American Journal of Agricultural Economics*, 95(2), 384–389.
- Menapace, L., Colson, G. & Raffaelli, R. (2016). A comparison of hypothetical risk attitude elicitation instruments for explaining farmer crop insurance purchases. *European Review of Agricultural Economics*, 43(1), 113–135.
- Moffatt, P. G. (2015). *Experimetrics: Econometrics for experimental economics*. Macmillan International Higher Education.
- Moffatt, P. G. (2021). Experimetrics: A survey. *Foundations and Trends in Econometrics*, 11(1–2), 1–152.
- Mograbi, E. (2022). Decision-makers are more impulsive on smartphones than on computers. *Journal of Behavioral and Experimental Economics*, 100, 101916.
- Morio, J., Balesdent, M., Jacquemart, D. & Vergé, C. (2014). A survey of rare event simulation methods for static input-output models. *Simulation Modelling Practice and Theory*, 49, 287–304.
- Nguyen, Q. (2010). How nurture can shape preferences: An experimental study on risk preferences of vietnamese fishers. *Environment and Development Economics*, 15(5), 609–631.

- Nguyen, Q. (2011). Does nurture matter: Theory and experimental investigation on the effect of working environment on risk and time preferences. *Journal of Risk and Uncertainty*, 43(3), 245–270.
- Parslow, E., Ranehill, E., Zethraeus, N., Blomberg, L., Schoultz, B., Hirschberg, A. L., Johannesson, M. & Dreber, A. (2019). The digit ratio (2D:4D) and economic preferences: No robust associations in a sample of 330 women. *Journal of the Economic Science Association*, 5(2), 149–169.
- Prelec, D. (1998). The probability weighting function. *Econometrica*, 66(3), 497–527.
- Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization*, 3(4), 323–343.
- Rabin, M. (1998). Psychology and economics. *Journal of Economic Literature*, 36(1), 11–46.
- Reynaud, A. & Couture, S. (2012). Stability of risk preference measures: Results from a field experiment on French farmers. *Theory and Decision*, 73(2), 203–221.
- Riddel, M. (2012). Comparing risk preferences over financial and environmental lotteries. *Journal of Risk and Uncertainty*, 45(2), 135–157.
- Riddel, M. & Kolstoe, S. (2013). Heterogeneity in life-duration preferences: Are risky recreationists really more risk loving?. *Journal of Risk and Uncertainty*, 46(2), 191–213.
- Rodriguez-Lara, I & Ponti, G. (2017). Social motives vs social influence: An experiment on interdependent time preferences. *Games and Economic Behavior*, 105, 177–194.
- Rubinstein, R. Y. and Kroese, D. P. (2017). *Simulation and the monte carlo method*. (3rd ed.) John Wiley & Sons, Inc.
- Saha, A. (1993). Expo-power utility: A ‘Flexible’ form for absolute and relative risk aversion. *American Journal of Agricultural Economics*, 75(4), 905–913.
- Santos-Pinto, L., Bruhin, A., Mata, J. & Åstebro, T. (2015). Detecting heterogeneous risk attitudes with mixed gambles. *Theory and Decision*, 79(4), 573–600.
- Shin, S., Magnan, N., Mullally, C. & Janzen, S. (2022). Demand for weather index insurance among smallholder farmers under prospect theory. *Journal of Economic Behavior & Organization*, 202, 82–104.
- Sprenger, C. (2015). An endowment effect for risk: Experimental tests of stochastic reference points. *Journal of Political Economy*, 123(6), 1456–1499.
- Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature*, 38(2), 332–382.
- Stern, S. (1997). Simulation-based estimation. *Journal of Economic Literature*, 35(4), 2006–2039.
- Stewart, M. B. (1983). On least squares estimation when the dependent variable is grouped. *The Review of Economic Studies*, 50(4), 737–753.
- Sugden, R., Zheng, J. & Zizzo, D. J. (2013). Not all anchors are created equal. *Journal of Economic Psychology*, 39, 21–31.
- Sun, T. and Ding, Y. (2019). *Copula-based semiparametric regression method for bivariate data under general interval censoring*. Biostatistics.
- Tanaka, T., Camerer, C. F. & Nguyen, Q. (2010). Risk and time preferences: Linking experimental and household survey data from Vietnam. *American Economic Review*, 100(1), 557–571.
- Tanaka, Y. & Munro, A. (2013). Regional variation in risk and time preferences: Evidence from a Large-scale field experiment in rural Uganda. *Journal of African Economics*, 23(1), 151–187.
- Train, K. (2009). *Discrete choice methods with simulation*. (2nd ed) Cambridge, UK, Cambridge University Press.
- Tversky, A. & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297–323.
- Veider, F. M. (2018). Violence and risk preference: Experimental evidence from Afghanistan: Comment. *American Economic Review*, 108(8), 2366–2382.
- Veider, F. M., Martinsson, P., Nam, P. K. & Truong, N. (2019). Risk preferences and development revisited. *Theory and Decision*, 86(1), 1–21.
- Vollmer, E., Hermann, D. and Musshoff, O. (2018) The disposition effect in farmers’ selling behavior – an experimental investigation. *Agricultural Economics*, 50.
- von Gaudecker, H. -M., van Soest, A. & Wengstrom, E. (2011). Heterogeneity in risky choice behavior in a broad population. *American Economic Review*, 101(2), 664–694.
- Voors, M. J., Nillesen, E. E. M., Verwimp, P., Bulte, E. H., Lensink, R. & Van Soest, D. P. (2012). Violent conflict and behavior: A field experiment in Burundi. *American Economic Review*, 102(2), 941–964.
- Wegmann, J., Hermann, D. and Musshoff, O. (2021) Assessing risk attitudes and time preferences on the rural–urban interface of Bengaluru, India. *Agricultural Finance Review*, ahead-of-print.
- Wilcox, N. T. (2017). Random expected utility and certainty equivalents: Mimicry of probability weighting functions. *Journal of the Economic Science Association*, 3(2), 161–173.
- Wilcox, N. T. (2023). Unusual estimates of probability weighting functions. In *Models of risk preferences: Descriptive and normative challenges*. (69–106), Emerald Publishing Limited.
- Yu, C. W., Zhang, Y. J. & Zuo, S. X. (2021). Multiple switching and data quality in the multiple price list. *The Review of Economics and Statistics*, 103(1), 136–150.

- Zhu, J., Shi, Y., Li, J. & Zhang, Z. (2020). Role of risk attitude and time preference in preventive aspirin use adherence. *Journal of Evaluation in Clinical Practice*, 26(3), 819–825.
- Zieglmeyer, F. & Zieglmeyer, M. (2014). Parenting is risky business: Parental risk attitudes in small stakes decisions on behalf of their children. *Review of Economics of the Household*, 14(3), 599–623.

Cite this article: Conte, A., Moffatt, P. G., & Riddell, M. (2026). The Multivariate Heterogeneous Preference estimator for Switching Multiple Price List data. *Experimental Economics*, 1–25. <https://doi.org/10.1017/eec.2026.10046>