

School of Economics Working Paper
2019-04



**SCHOOL OF
ECONOMICS**

**The Multivariate Random Preference
Estimator for Switching Multiple Price
List Data**

by Anna Conte *
Peter G. Moffatt**
Mary Riddel***

*Sapienza University of Rome

**University of East Anglia

***University of Nevada

School of Economics
University of East Anglia
Norwich Research Park
Norwich NR4 7TJ
United Kingdom
www.uea.ac.uk/economics

The Multivariate Random Preference Estimator for Switching Multiple Price List Data

Anna Conte^{a,*}, Peter G. Moffatt^{b,†}, Mary Riddell^{c,‡}

^a*Department of Statistical Sciences, Sapienza University of Rome, Italy*

^b*School of Economics, University of East Anglia, Norwich, UK*

^c*Department of Business, University of Nevada, Las Vegas, USA*

Abstract

The use of Multiple Price Lists to elicit individuals' risk preferences is widespread. To model data collected through this method, we introduce the Multivariate Random Preference (MRP) estimator, specifically designed for the “switching” variant of such lists. This is a new estimation approach that enables us to exploit all available information derived from subjects' switch points in the lists. Monte Carlo simulations show that our estimator is consistent and has good small-sample properties. The estimator is derived for a two-parameter model in a risky context.

Keywords: Risk Preference; Monte Carlo Simulations; Importance Sampling

JEL Classification: C51; C52; C91; D81

**Email:* anna.conte@uniroma1.it

†*Email:* p.moffatt@uea.ac.uk

‡*Email:* mary.riddell@unlv.edu

1 Introduction

Multiple Price Lists (MPL) are widely used in Experimental Economics to elicit individuals' preferences. In the time since the MPL was first introduced by Holt and Laury (2002) to measure attitude to risk in a laboratory context, the technique has steadily gained popularity. Its use is encroaching upon field experiments associated or not with sample surveys, whenever some measures of individual preferences may be considered as suitable predictors for some phenomena.¹

Andersen et al. (2006) discuss and explain the basic MPL technique and its variants. Apropos they write: "In the context of eliciting a willingness to pay for some commodity, it confronts the subject with an array of ordered prices in a table, one per row, and asks the subject to indicate "yes" or "no" for each price. The experimenter then selects one row at random, and the subject's choice for that row is implemented".

The Switching Multiple Price List (sMPL) varies the basic design by asking the subject to state the row she wants to switch in. Andersen et al. (2006) argue that sMPL has additional advantages over simple lotteries and certainty equivalent methods because it is comparatively simple for subjects to master and it forces monotonicity in preferences. The approach is described in detail also in Tanaka et al. (2010), Riddell (2012) and Riddell and Kolstoe (2013).

Data collected via MPL and sMPL can then be used to estimate parameters of interest which characterise the relevant choice model. Estimating a one-parameter model from MPL data is rather straightforward, but the case of multiple parameters is not so. In this methodological paper, we develop the Multivariate Random Preference (MRP) estimator for data from sMPL, specifically designed to estimate preference functions characterised by more than one parameter. We use as a working example the case of the estimation of a two-parameter preference function in a risky context, but our estimator is prone to be extended to any number of parameters and any context. By means of a Monte Carlo simulation, we assess the properties of our estimator.

The paper develops as follows. Section 2 explains the Random Preference choice model and how MPL data is used to fit it when preference functions are characterised by one single parameter. Section 3 extends the Random Preference choice model to multiple-parameter preference functions and introduces the sMPL approach as a tool to collect data to estimate it, by overcoming some drawbacks of MPL series. In Section 4, we explain how sMPL series are used to delimit the areas where parameters are located when multi-parameter preference functions are assumed. Section

¹The number of articles to refer to is enormous, and it is outside the main scope of this work that of reviewing them. Here, we cite as examples of the use of MPL data the experiment by Bruhin et al. (2010), the field experiment by Harrison et al. (2002) and the field experiment administered to a representative sample of a population by von Gaudecker et al. (2011).

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	$\lambda < -1.72$
2	2/10 of \$2, 8/10 of \$1.60	2/10 of \$3.85, 8/10 of \$0.10	\$0.83	$-1.72 < \lambda < -0.95$
3	3/10 of \$2, 7/10 of \$1.60	3/10 of \$3.85, 7/10 of \$0.10	\$0.50	$-0.95 < \lambda < -0.49$
4	4/10 of \$2, 6/10 of \$1.60	4/10 of \$3.85, 6/10 of \$0.10	\$0.16	$-0.49 < \lambda < -0.15$
5	5/10 of \$2, 5/10 of \$1.60	5/10 of \$3.85, 5/10 of \$0.10	-\$0.18	$-0.15 < \lambda < 0.15$
6	6/10 of \$2, 4/10 of \$1.60	6/10 of \$3.85, 4/10 of \$0.10	-\$0.51	$0.15 < \lambda < 0.41$
7	7/10 of \$2, 3/10 of \$1.60	7/10 of \$3.85, 3/10 of \$0.10	-\$0.85	$0.41 < \lambda < 0.68$
8	8/10 of \$2, 2/10 of \$1.60	8/10 of \$3.85, 2/10 of \$0.10	-\$1.18	$0.68 < \lambda < 0.97$
9	9/10 of \$2, 1/10 of \$1.60	9/10 of \$3.85, 1/10 of \$0.10	-\$1.52	$0.97 < \lambda < 1.37$
10	10/10 of \$2, 0/10 of \$1.60	10/10 of \$3.85, 0/10 of \$0.10	-\$1.85	$\lambda > 1.37$

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

5 describes the estimation approach based on midpoints adopted by other researchers and its drawbacks. Section 6 develops the MPL estimator using Importance Sampling, a Monte Carlo simulation technique for drawing from densities. Section 7 reports from a Monte Carlo simulation technique which proves that the MRP estimator is consistent and performs well in small samples, while the midpoint estimator is not. Section 8 concludes.

2 The basic MPL design and the Random Preference model

One of the most used stochastic approaches to modelling individuals' choices is the Random Preference (RP) choice model, introduced by Loomes and Sugden (1995). In brief, the idea behind it is that each subject draws a value for the parameters characterising some utility function from a distribution, and the decision is made on their basis. Choices occur without error.²

Estimating this choice model means estimating that distribution. For doing this, data is commonly collected from a sample via MPL series, as, in a risky context, in the notorious article by Holt and Laury. There experimental subjects are presented with the one list of ten rows reported in Table 1, each displaying two alternative two-prize lotteries, A and B. In each row, the probability of the prizes of the otherwise-equal lotteries varies such that one of the two (lottery B, the riskiest in the table) appears more and more appealing. Subjects are asked to indicate their preferred lottery in each row.

Subjects with monotonic preferences are expected to choose option A (the safer lottery) up to a certain row to switch to lottery B in all the subsequent rows. We will refer to the row where the subject switches from one lottery to the other as the “switch point”.

²For more details on the RP model and its differences with the Random Utility model, see Conte and Hey (2018) and Conte and Moffatt (2014). The latter also discuss the Random Behavioural choice model.

Under the assumption that subjects are Expected Utility (EU) maximisers, given a utility function, the subject’s switch point implies for that subject a range of values of the parameter of the chosen utility function.³ In the list at hand, a Constant Relative Risk Aversion (CRRA) utility function with specification $U(x) = x^{1-\lambda}/(1-\lambda)$ would produce the ranges reported in Table 1. In this one-parameter utility function, λ is the relative risk aversion parameter, taking negative (positive) values for risk lovers (aversers) and equalling zero for risk neutrals. In the list, later switch points indicate higher levels of risk aversion.

Given the choices of switch points from a sample, which define the range of values for λ for each subject in the sample, the distribution of λ can be estimated by simple interval regression, as explained in more detail in Appendix A.1.

All in all, the RP model is easy in every respect: MPL choice tasks are straightforward, quick and easy to introduce; 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 (the interval regression) is available in the most popular statistical packages.⁴ However, there are at least three cases, discussed below, that render such a procedure difficult or unsuitable.

First of all, it is noted that often subjects switch more than once when scrolling through the list. Already Holt and Laury (2002) report of roughly 10% of such inconsistencies. Additional error stories to the basic RP model are needed to capture this behaviour. Typically, researchers add a “tremble” to the RP choice model, which capture the possibility that subjects from time to time loose concentration and choose at random. The tremble may or may not be combined with the hypothesis that a new value of λ is drawn in each row, and the decision as to whether to choose which lottery is made on its basis.⁵ Some others resort to the Random Utility model. However, we strongly believe that the assumption of (conditional) independence underlying that stochastic model does not hold, when the tasks in the MPL are not presented one at a time in a random order but altogether as in Table 1.⁶ All of these are certainly noteworthy stories, though not strongly related to this work.⁷

³We will get back to the way of calculating the ranges later on. For more details, see Appendix A.

⁴The standard interval regression can be promptly applied to MPL-derived ranges of a parameter whenever its distribution can be assumed to be normal or lognormal, because in-built in mostly all the most popular statistical packages. For different distributions than these, programming the likelihood function does not require particular skills.

⁵A tremble can also help explain the cases in which a stochastically dominated alternative is chosen (see Loomes et al. 2002).

⁶We have to acknowledge that Andersen et al. (2006) write that their “findings indicate that nothing is lost from using an enforced single switching point, as in the sMPL design” and conclude that “the switching behavior that is often observed in MPL simply reflects indifference”.

⁷For other issues related to the use of MPL designs, see Bosch-Domènech and Silvestre (2013) and Freeman et al. (2019).

Secondly, there exist a huge literature claiming that the EU hypothesis cannot accommodate the vast majority of people’s choices under risk. Different models than EU, usually more extravagant in terms of parameters, are required.

Finally, whenever the simple RP model cannot be used, interval regression is not applicable as is. In this work, we specifically deal with the estimation of the multiple-parameter RP model, which requires a not-so-straightforward extension of the interval regression approach.

3 The sMPL design and the Multivariate Random Preference model

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* preferences to be monotonic.

Table 2 shows two examples of sMPL series first used in Tanaka et al. (2010) and then in Riddell (2012) and Riddell and Kolstoe (2013). As before, the 14-row lists display two alternative lotteries. In this case, in each list, lottery A, all probabilities and one of lottery B’s prizes are kept fix. What changes is the largest prize of lottery B which increases through rows, making it more and more appealing in terms of expected payoff.⁸

Essentially, the sMPL design rules out the complications related to non-monotonicity of preferences, enabling us to focus on the application and estimation of the RP choice model to the case of more than one parameter.⁹

As noted earlier, there is overwhelming empirical evidence that the EU model is unable to explain the most part of individuals’ behaviour under risk. Numerous studies have shown that the majority of people tend to overemphasize favorable unlikely events and deemphasize favorable likely events in their financial decision making relative to the predictions of EU theory (among others, Tversky and Kahneman (1992), Quiggin (1982), Rabin (1998), Starmer (2000), Bruhin et al. (2010), Conte et al. (2011)). Cumulative Prospect theory and Rank Dependent Utility (RDU) theory allow for violations of the independence axiom by transforming objective probabilities into decision weights using a probability-weighting function. This helps explain why people, for

⁸We note that the difference between MPL and sMPL is not in the way lotteries differ across rows (by changing probabilities and/or prizes), but in the rules given to subjects concerning the way to proceed to the subsequent row after having made a choice.

⁹In any way this has to be intended as a claim that sMPL is a superior tool with respect to MPL to measure subjects’ risk preference. Whether one or the other is more up to that target is a methodological issue proper to experimentalists. Nevertheless, finding a good-property multivariate estimator for data from sMPL is an interesting exercise.

TABLE 2
THE PAIRED LOTTERY CHOICES USED IN TANAKA ET AL. (2010)

<i>LIST 1</i>									
Row	Lottery A				Lottery B				expected payoff difference
	probability (\$400)	payoff	probability (\$100)	payoff	probability (\$?)	payoff	probability (\$50)	payoff	
1	0.3	\$400	0.7	\$100	0.1	\$680	0.9	\$50	\$77
2	0.3	\$400	0.7	\$100	0.1	\$750	0.9	\$50	\$70
3	0.3	\$400	0.7	\$100	0.1	\$830	0.9	\$50	\$62
4	0.3	\$400	0.7	\$100	0.1	\$930	0.9	\$50	\$52
5	0.3	\$400	0.7	\$100	0.1	\$1,060	0.9	\$50	\$39
6	0.3	\$400	0.7	\$100	0.1	\$1,250	0.9	\$50	\$20
7	0.3	\$400	0.7	\$100	0.1	\$1,500	0.9	\$50	-\$5
8	0.3	\$400	0.7	\$100	0.1	\$1,850	0.9	\$50	-\$40
9	0.3	\$400	0.7	\$100	0.1	\$2,200	0.9	\$50	-\$75
10	0.3	\$400	0.7	\$100	0.1	\$3,000	0.9	\$50	-\$155
11	0.3	\$400	0.7	\$100	0.1	\$4,000	0.9	\$50	-\$255
12	0.3	\$400	0.7	\$100	0.1	\$6,000	0.9	\$50	-\$455
13	0.3	\$400	0.7	\$100	0.1	\$10,000	0.9	\$50	-\$855
14	0.3	\$400	0.7	\$100	0.1	\$17,000	0.9	\$50	-\$1555

<i>LIST 2</i>									
Row	Lottery A				Lottery B				expected payoff difference
	probability (\$400)	payoff	probability (\$300)	payoff	probability (\$?)	payoff	probability (\$50)	payoff	
1	0.9	\$400	0.1	\$300	0.7	\$540	0.3	\$50	-\$3
2	0.9	\$400	0.1	\$300	0.7	\$560	0.3	\$50	-\$17
3	0.9	\$400	0.1	\$300	0.7	\$580	0.3	\$50	-\$31
4	0.9	\$400	0.1	\$300	0.7	\$600	0.3	\$50	-\$45
5	0.9	\$400	0.1	\$300	0.7	\$620	0.3	\$50	-\$59
6	0.9	\$400	0.1	\$300	0.7	\$650	0.3	\$50	-\$80
7	0.9	\$400	0.1	\$300	0.7	\$680	0.3	\$50	-\$101
8	0.9	\$400	0.1	\$300	0.7	\$720	0.3	\$50	-\$129
9	0.9	\$400	0.1	\$300	0.7	\$770	0.3	\$50	-\$164
10	0.9	\$400	0.1	\$300	0.7	\$830	0.3	\$50	-\$206
11	0.9	\$400	0.1	\$300	0.7	\$900	0.3	\$50	-\$255
12	0.9	\$400	0.1	\$300	0.7	\$1,000	0.3	\$50	-\$325
13	0.9	\$400	0.1	\$300	0.7	\$1,100	0.3	\$50	-\$395
14	0.9	\$400	0.1	\$300	0.7	\$1,300	0.3	\$50	-\$535

Note: In each list, subjects are asked to start choosing between lottery A and B from the first row, and to proceed with the following rows unless lottery B is chosen.

example, buy lottery tickets and simultaneously over-insure against financial losses. That is a behavioural pattern that the EU model is unable to explain.

The simplest probability weighting function used to deal with such probability distortions is characterised by at least one parameter, which adds to the coefficient of risk attitude to be estimated. Therefore, when an RDU hypothesis is made, each subject is thought to draw a pair of values – one for the risk attitude and one for the probability distortion coefficient – from their joint distribution and make choices over the sMPL series based on those values. In this case the RP choice model requires a multivariate estimator in order to estimate the joint distribution of such coefficients.

4 Use of sMPL and the parameter-areas

In this section, we show how to extract from sMPL series the information required to estimate the RP choice model. For this purpose, we continue with the example of a two-parameter model for risky choice data.

Let us start by noting that the problem of simultaneous estimation of risk attitude and probability weighting coefficients requires for each subject to collect choices from as many *ad hoc* sMPL series as the number of parameters to be estimated. The reason for this will become clear later. For the purpose of this exercise, we will use the two lists in Table 2, which we will denote as List 1 ($L = 1$) and List 2 ($L = 2$).

Therefore, subject i , $i = 1, \dots, n$, faces two price lists, $L = 1, 2$. Each list comprises $T - 1$ rows indexed by t .¹⁰ Each row displays two lotteries, lottery A^L and lottery $B^{L,t}$, $t = 1, \dots, T$, where $t = T$ indicates that the subject never switches to lottery B by always preferring lottery A.

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.

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

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

Here, $u_i(z)$ is a utility function, where z is the lottery outcome, and $w_i(r)$ is a probability-weighting function, where r is the true probability.

For the utility function, we adopt the CRRA-power functional form, $u_i(z) = z^{\alpha_i}$, where z is the prize. The coefficient $\alpha_i > 0$ is less than 1 for risk-averse agents, equal to 1 for risk-neutral agents, and greater than 1 for risk-loving agents.

The probability weighting function is assumed to take on the one-parameter functional form proposed by Prelec (1998), $w_i(r) = \exp[-(-\ln(r))^{\gamma_i}]$. 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; as $\gamma_i \rightarrow 0$, $w_i(r)$ becomes a step function, that is flat everywhere except the endpoints of

¹⁰The lists need not to contain the same number of rows. Here, for simplicity, we use a notation which strictly refers to the working case, but its generalisation is straightforward.

the probability interval; when $0 < \gamma_i < 1$, the probability weighting function takes on an inverse-s shape. In this case, subjects overemphasise low-probability high outcomes in their decision making relative to EU theory. When $\gamma_i > 1$, the probability weighting function takes on an s-shape, so that, relatively to an EU decision maker, 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 have to satisfy the following conditions, denoted as $C_i^L(\tilde{t}^L)$ or, more briefly, as C_i^L .

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

These conditions, applied to List 1 and List 2, produce the plots in Figure 1, top-left panel and top-right panel, respectively.¹¹ 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 switched in row 1, then such pairs of values would be all those above the highest line. 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 lines. This reasoning continues up to the lowest line, which instead delimits from above all the values compatible with having never switched to lottery B. Conditions (2) when applied to lottery choices in List 1 define the upward sloping lines, when applied to lottery choices in List 2 define the downward sloping lines. In both lists, the higher the lines the earlier in the sequence the subject switches to lottery B.

How insightful are these areas? Let us consider a subject who switches to lottery B in row 7 of List 1. The pairs of values compatible with that choice are comprised between the 6th and 7th curve from above. 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 other switch points of the two lists.¹² 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

¹¹ *Mutatis mutandis*, these same conditions apply to the list in Table 1 to compute the ranges of λ . In that case, $w_i(r) = r$ such that the EU hypothesis substitutes the RDU's. Further details can be found in Appendix A.

¹²Crucially, 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.

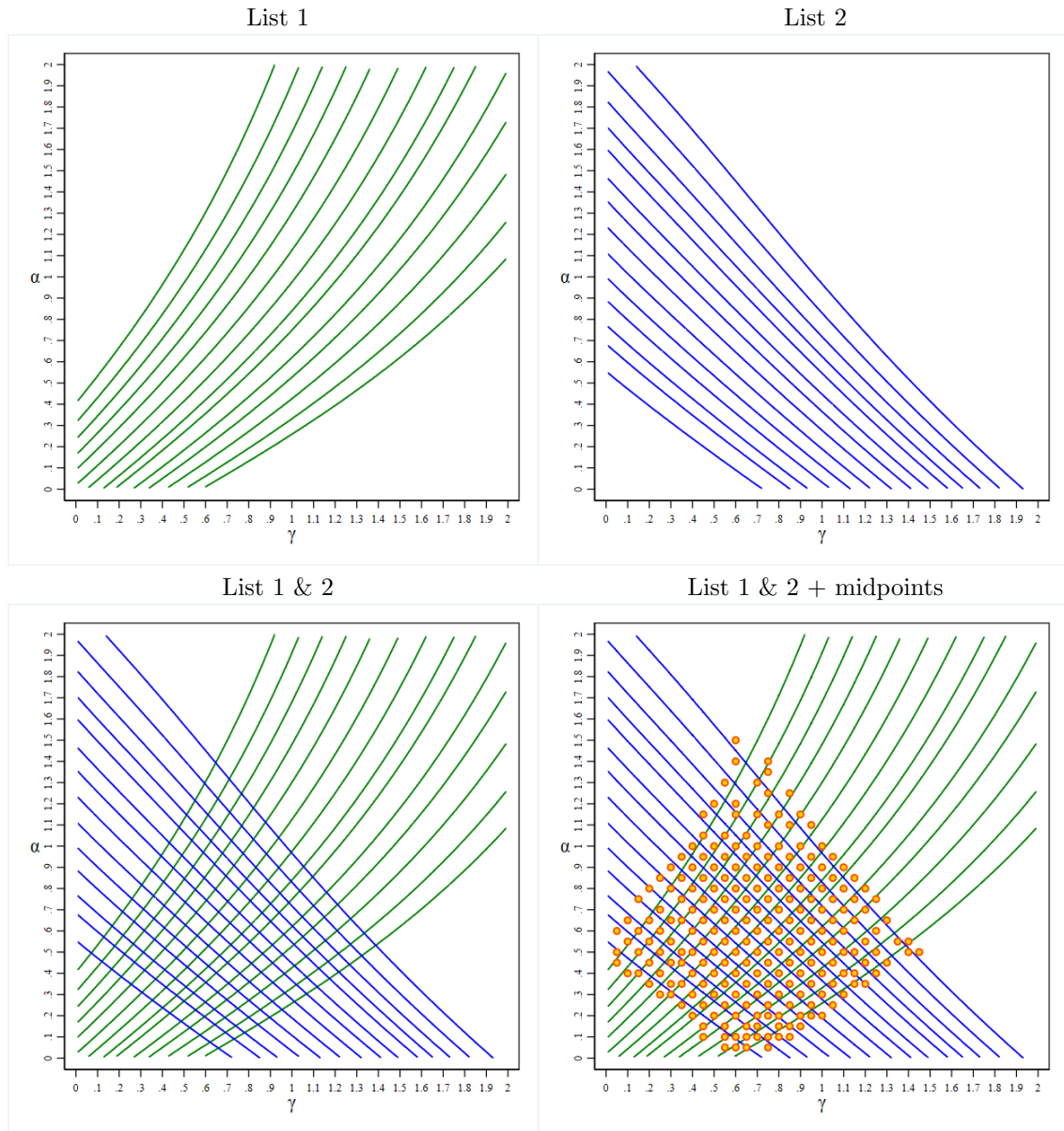


FIGURE 1
RANGES FOR α AND γ IMPLIED BY CHOICES OVER THE LOTTERIES IN TABLE 2

circumscribe more tightly the area containing the values of the coefficients which apply to each type of behaviour under risk.

In effect, with the exception of the extreme choices (that are that of 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 1, bottom-left panel.

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

5 Estimation using midpoints

The problem of estimating data in the irregular areas obtained as explained above has already been considered by Tanaka et al. (2010), Liu (2013), Riddell (2012) and Riddell and Kolstoe (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), represented by yellow dots, are superimposed to the (α, γ) -areas in the bottom-right panel of Figure 1. The cartesian coordinates of these midpoints (abscissae for γ and ordinates for α) are then used to fit linear models which estimate the distribution of each parameter in order to investigate its variation arising from demographic or other variables.

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 switches in row 1 and 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 falls 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 parameters α and γ are not estimated jointly, so a significant amount of information is lost and the precision of the parameter estimates is compromised.¹³ Figure 1 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 coefficients are modelled jointly. Therefore, the information provided by this graph concerning location and dimension of each area is crucial in order to make inference.

¹³Often, 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 is assumed a distribution with a domain which extends to the non-positive reals.

The approach detailed below addresses all these issues, since it jointly estimates parameters 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 these midpoints, that we use as a countercheck of the validity of our estimator through a Monte Carlo simulation.

6 The Multivariate Random Preference estimator with Importance Sampling

As explained, to be able to estimate the joint distribution of risk attitude and probability weighting, information is needed 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 its bivariate case. It essentially consists of a generalization of the interval regression model to two dimensions and 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 Random Preference (MRP) estimator for sMPL data.

The two parameters of interest are described by a joint distribution with moments which may be allowed to vary by subject's characteristics. Since α_i and γ_i are both larger than 0, for the purpose of this exercise, we assume that they follow a bivariate lognormal distribution¹⁴

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

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

¹⁴Alternatively, 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.

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

where $\mathbf{1}[\cdot]$ is an indicator function taking the value 1 if the statement in brackets is true, 0 otherwise. It guarantees that conditions (2) for List 1 and 2 are jointly satisfied, that is what is required to identify the (α, γ) -area corresponding to i 's switch points.

This two-dimensional integral has no closed form solution, regardless of the distributional assumptions relative to α and γ .¹⁵ 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.¹⁶

After having unsuccessfully reviewed the most obvious solutions, to calculate probability (4), we resorted to Monte Carlo integration, which raises a number of issues anyhow. First of all, let us consider the problem of sampling from any of the irregularly-shaped areas in Figure 1, bottom-left panel. They are all extremely small. 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,¹⁷ which is the result of the simplest and most widely used sampling procedures, this would not be guaranteed even by allowing a number of simulated draws per subject in the order of billions.¹⁸ This is due to the so-called “rare event sampling” problem, which has been largely studied 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, the simulator is a step function in α and γ . This is due to the presence of the indicator function in (4). This problem has already been extensively analysed and addressed for when simulations are used to estimate discrete choice models in McFadden (1984) and McFadden and Ruud (1994). For an extensive review, see Stern (1997), who writes: “Step functions are very

¹⁵The dimension of the integral in the probability density function increases with the number of parameters to be estimated.

¹⁶Although the bivariate interval regression is not readily available from statistical packages, it is not difficult to program for a number of popular bivariate distributions. More on this in Appendix A.2.1.

¹⁷This “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).

¹⁸More details on this are available from the authors on request.

difficult to optimize over because derivatives are equal to zero everywhere but points with discrete steps and do not exist at such points”.

Finally, the variance of the estimator is “unnecessarily large” (Stern 1997).

Upon considering many other simulation techniques for drawing from a multivariate density (the Acceptance-Rejection, the Gibbs sampler, the Metropolis-Hastings, and so on), our choice has descended on using the Importance Sampling technique.¹⁹ This is because of the appreciable similarity between the probability in (4) and that of the Multinomial Probit model (see Stern (1997) [pag. 2013, Eq. (2.21)]), which would have 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 (which do not commonly affect the Multinomial Probit estimator, though) can be easily treated.

To see how this technique operates, consider rewriting the probability of switching in row h of List 1 and in row k of List 2 as

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

Here, we have innocuously divided and multiplied the integrand by $g(\alpha, \gamma)$, which is a distribution in α and γ easy to sample from and that is concentrated in an area where we know that the values of α and γ lie for subjects i . This is the basic trick for drawing from densities referred to as “Importance Sampling”,²⁰ 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” $f(\alpha, \gamma) / g(\alpha, \gamma)$.

In our working case, the “important” part is suggested by the grid in Figure 1, bottom-left panel, which tells us where the combinations of α_i and γ_i are for a given choice of switch points in the two lists. Essentially, a judicious choice of $g(\alpha, \gamma)$ is all that we need for the MRP estimator to work.

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 2 magnifies the irregularly-shaped area from Figure 1, bottom-left plot, that encloses all combinations of α and γ compatible with that choice. The red lines delineate a rectangle around that area which exactly contains it, establishing a lower

¹⁹We 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 reside in the way of drawing from such distributions.

²⁰For a detailed description, see Gouriéroux and Monfort 1996, among others.

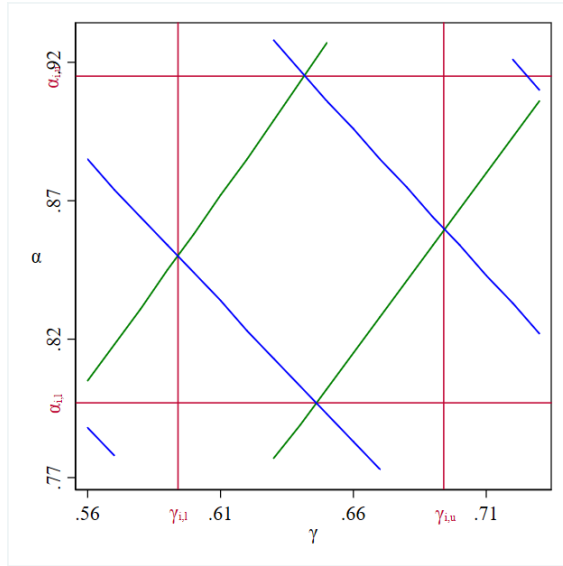


FIGURE 2
 (α, γ) -RANGE FOR SWITCHING IN ROW 4 OF LIST 1 AND ROW 5 OF LIST 2

and upper limit for the parameter α , that is $\alpha_{i,l}$ and $\alpha_{i,u}$, and γ , that is $\gamma_{i,l}$ and $\gamma_{i,u}$, respectively. In order to circumvent the rare-event sampling issue discussed earlier, we can choose $g(\alpha, \gamma)$ so that it has exactly that rectangle as a support. Any bivariate distribution truncated within the so-established limits may work, and it is even possible to choose two different and independent distributions for the two dimensions. This makes sure that we will certainly hit the (α, γ) -area inscribed in the rectangle from which we ultimately want to sample.²¹

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 what the “importance weight” $f(\alpha, \gamma) / g(\alpha, \gamma)$ accomplishes in Eq. (5).

6.1 The steps of Importance Sampling

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

- i. 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 want to sample from;

²¹This is not necessarily true for the open intervals. In those cases, the rectangles which frame the relevant (α, γ) -area are only limited from below. 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.

- ii. compute the indicator $\mathbf{1} [C_i^1 (h; \hat{\alpha}_r, \hat{\gamma}_r) \& C_i^2 (k; \hat{\alpha}_r, \hat{\gamma}_r)]$ following Eqq. (1) and (2);
- iii. multiply by the importance weight $f(\hat{\alpha}_r, \hat{\gamma}_r) / g(\hat{\alpha}_r, \hat{\gamma}_r)$;
- iv. 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:

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

which for $R \rightarrow \infty$ approximates the integral (5) with an infinite sum. This is the individual likelihood contribution that is passed to the likelihood function to maximise.²² The so-obtained MRP estimator incorporates all 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.

Importance Sampling estimators are proven to be consistent, smooth and to have a reduced variance. For an explanation and extensive discussion of all the properties of this type of estimators see Stern (1997).

7 A Monte Carlo simulation study

In this section, we investigate the properties of the MRP estimator via a Monte Carlo simulation. The performance of the estimator is compared with that of the midpoint estimator improved so that it estimates the *joint* distribution of α and γ . Details on this modified estimator can be found in Appendix B.

We used 4 Data Generating Processes (DGP) with parameters corresponding to the bivariate lognormal distributions shown in Figure 3 and reported in column 2 of Table 3. DGP1-2 are chosen so as to represent distributions compatible with the most common empirical findings (the first is more dispersed, the latter more concentrated around the parameter means). DGP3, instead, determines a symmetric distribution very much concentrated towards small parameter values. The reason for exploring the performance of the MRP estimator with this distribution is that of verifying whether the two lists may by themselves introduce some bias in the direction of one or the other parameter together with verifying whether it works well with such an extreme distribution mostly generated from switching later in the sequence in both lists. DGP4 depicts a distribution

²²The program is written in Stata 15.

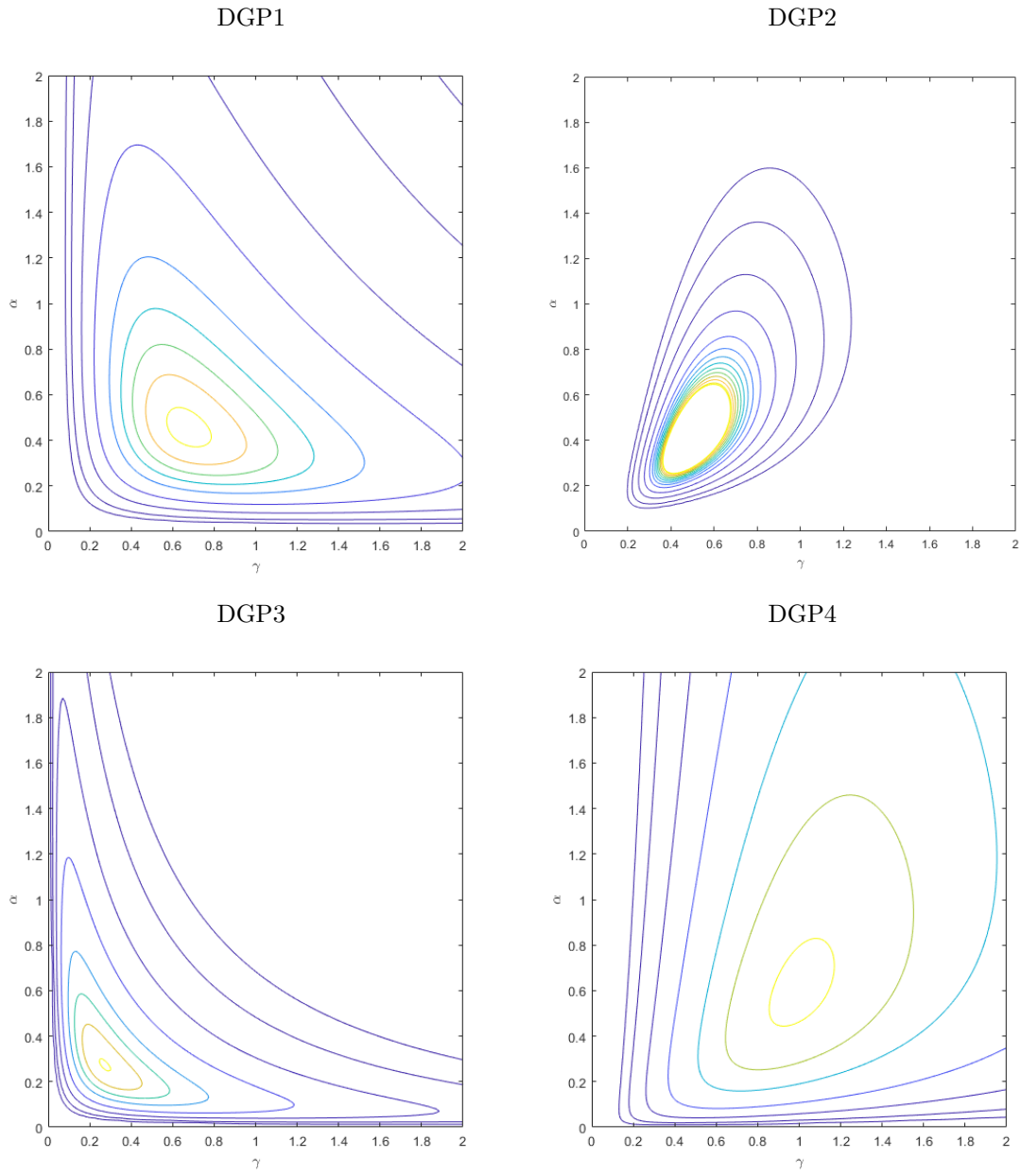


FIGURE 3
PLOTS OF THE BIVARIATE LOGNORMAL DISTRIBUTIONS USED IN THE SIMULATION

very much dispersed, which produces pairs of parameter values very much concentrated on the open intervals determined by early switch points.

From each distribution, 5000 simulated samples of 25, 100 and 500 observations each are drawn. Each observation corresponds to a choice of two switch points, one from List 1 and one from List 2 for a given pair of draws of α and γ .

To evaluate the performance of the two estimators, for each set of DGP and each sample size, we report the mean of the parameter estimates, the absolute bias as a measure of bias, the root mean square error (rMSE) as a measure of accuracy, and, as a measure of coverage, we use the actual size (the “size” is the probability to reject the null hypothesis when it is true) calculated as the proportion of times the 95% confidence interval does not contain the true value of the parameter, the nominal size being 0.05.

As a sampling distribution $g(\alpha, \gamma)$, we use two independent normal distributions with mean 0.5 and standard deviation 1, appropriately truncated within the limits $(\alpha_{i,l}, \alpha_{i,u})$ for α and $(\gamma_{i,l}, \gamma_{i,u})$ for γ , calculated as explained for each pair of switch points.

To draw from $g(\alpha, \gamma)$, we use two sequences of $R = 100$ Halton draws per observation (Train 2009).

The simulation results are displayed in Table 3. Let us concentrate on DGP1-3 first. In these cases, the MRP estimator shows a small bias in all the parameter estimates in the 25-observation samples which fades away when the sample size increases. A similar pattern is shown by the accuracy with noticeably decreasing root mean square errors. All this confirms the consistency of the estimator and its good small-sample properties. The actual size is slightly larger than 0.05 in the smallest samples, but it already approaches that figure for the 100-observation samples.

In stark contrast, the improved midpoint estimator shows a persistent bias in all the parameter estimates that does not decrease in larger samples and even increases sometimes. The root mean square errors decrease but not as convincingly as for the MRP estimator. This together with the large steadfast bias proves that the midpoint estimator is inconsistent. On top of that, the extremely high sizes, which tend to and even reach 1 especially in large samples, undoubtedly invalidate inference.²³ Sincerely, we would have bet that the lack of statistical properties of the midpoint estimator had to be primarily attributed to the observations corresponding to the open-ended intervals. Therefore, we would have expected to observe a better performance of that estimator for DGP2, which generates samples of observations mostly concentrated in the thick grid. Incidentally, in that case, two of the parameters (the standard deviations) show a smaller bias with respect to the MRP estimator. However, this is limited to the smallest sample size only, since their bias increases in the larger cases.

A special comment deserve the results from the last DGP. In this case, even the MRP estimator proves inconsistent, albeit its performance is still better than that of the midpoint estimator. The

²³Note that size is approaching 1 because incorrect values are being estimated more precisely.

TABLE 3
SIMULATION RESULTS FROM 5000 REPLICATIONS

MRP estimates - DGP1													
		25 obs.				100 obs.				500 obs.			
coefficient	true value	mean	bias	rMSE	size	mean	bias	rMSE	size	mean	bias	rMSE	size
μ_α	-.8	-.8024	.0024	.1273	.0570	-.8009	.0003	.0630	.0486	-.7999	.0001	.0287	.0536
σ_α	.6	.5902	.0098	.1044	.0904	.5984	.0016	.0504	.0596	.6005	.0005	.0227	.0542
μ_γ	-.4	-.4012	.0012	.1043	.0664	-.3997	.0003	.0515	.0562	-.3997	.0003	.0229	.0516
σ_γ	.5	.4918	.0082	.0807	.0802	.4982	.0018	.0401	.0568	.5002	.0002	.0179	.0518
ρ	-.4	-.3885	.0115	.2142	.1066	-.3913	.0087	.1074	.0670	-.3940	.0060	.0474	.0494
Midpoint estimates - DGP1													
		25 obs.				100 obs.				500 obs.			
coefficient	true value	mean	bias	rMSE	size	mean	bias	rMSE	size	mean	bias	rMSE	size
μ_α	-.8	-.7534	.0466	.1173	.0918	-.7535	.0465	.0706	.1342	-.7544	.0456	.0519	.4620
σ_α	.6	.5297	.0703	.1033	.2164	.5425	.0575	.0688	.3400	.5464	.0536	.0562	.8636
μ_γ	-.4	-.4116	.0116	.0913	.0730	-.4112	.0112	.0463	.0584	-.4114	.0114	.0229	.0858
σ_γ	.5	.4308	.0692	.0902	.2490	.4415	.0585	.0655	.4722	.4438	.0562	.0578	.9754
ρ	-.4	-.2111	.1889	.2399	.0888	-.2091	.1909	.2038	.5168	-.2071	.1929	.1954	.9998
MRP estimates - DGP2													
		25 obs.				100 obs.				500 obs.			
coefficient	true value	mean	bias	rMSE	size	mean	bias	rMSE	size	mean	bias	rMSE	size
μ_α	-.9	-.9017	.0017	.0618	.0686	-.9008	.0008	.0306	.0528	-.9004	.0004	.0134	.0490
σ_α	.3	.2923	.0077	.0461	.0862	.2988	.0012	.0227	.0606	.3006	.0006	.0101	.0512
μ_γ	-.7	-.6987	.0013	.0410	.0598	-.7004	.0005	.0207	.0586	-.7002	.0002	.0091	.0488
σ_γ	.2	.1946	.0054	.0313	.0892	.1993	.0007	.0154	.0600	.2003	.0003	.0068	.0486
ρ	.6	.5944	.0056	.1473	.1010	.6013	.0013	.0684	.0646	.6010	.0010	.0305	.0522
Midpoint estimates - DGP2													
		25 obs.				100 obs.				500 obs.			
coefficient	true value	mean	bias	rMSE	size	mean	bias	rMSE	size	mean	bias	rMSE	size
μ_α	-.9	-.8605	.0395	.0730	.1282	-.8607	.0393	.0497	.2610	-.8606	.0394	.0416	.8282
σ_α	.3	.2952	.0048	.0428	.0796	.3028	.0028	.0215	.0488	.3040	.0040	.0104	.0602
μ_γ	-.7	-.7207	.0207	.0461	.1008	-.7218	.0218	.0301	.1922	-.7216	.0216	.0235	.6454
σ_γ	.2	.2008	.0008	.0282	.0618	.2058	.0058	.0153	.0488	.2070	.0070	.0094	.1600
ρ	.6	.4878	.1121	.1927	.0814	.4931	.1069	.1302	.2416	.4927	.1073	.1123	.9176
MRP estimates - DGP3													
		25 obs.				100 obs.				500 obs.			
coefficient	true value	mean	bias	rMSE	size	mean	bias	rMSE	size	mean	bias	rMSE	size
μ_α	-1.3	-1.3073	.0073	.1585	.0556	-1.3038	.0038	.0774	.0566	-1.3027	.0027	.0337	.0484
σ_α	.7	.6849	.0150	.1262	.0966	.6977	.0023	.0586	.0556	.7001	.0001	.0261	.0478
μ_γ	-1.3	-1.3034	.0034	.1557	.0590	-1.3021	.0021	.0782	.0560	-1.3015	.0015	.0341	.0466
σ_γ	.7	.6816	.0184	.1268	.0962	.6963	.0037	.0627	.0644	.6994	.0006	.0272	.0512
ρ	-.7	-.6789	.0210	.1845	.1492	-.6862	.0138	.0863	.0666	-.6897	.0102	.0376	.0492
Midpoint estimates - DGP3													
		25 obs.				100 obs.				500 obs.			
coefficient	true value	mean	bias	rMSE	size	mean	bias	rMSE	size	mean	bias	rMSE	size
μ_α	-1.3	-1.1893	.1106	.1672	.1808	-1.1884	.1116	.1281	.4388	-1.1888	.1112	.1146	.9778
σ_α	.7	.6087	.0913	.1272	.2458	.6239	.0761	.0875	.4188	.6274	.0725	.0751	.9502
μ_γ	-1.3	-1.1913	.1086	.1672	.1764	-1.1920	.1080	.1257	.4102	-1.1924	.1076	.1112	.9676
σ_γ	.7	.6150	.0850	.1260	.2378	.6310	.0689	.0841	.3672	.6352	.0648	.0681	.8710
ρ	-.7	-.6297	.0703	.1509	.0766	-.6210	.0790	.1017	.1980	-.6186	.0814	.0863	.8614
MRP estimates - DGP4													
		25 obs.				100 obs.				500 obs.			
coefficient	true value	mean	bias	rMSE	size	mean	bias	rMSE	size	mean	bias	rMSE	size
μ_α	-.5	-.5651	.0651	.2021	.0636	-.5570	.0570	.1116	.0856	-.5530	.0530	.0673	.2240
σ_α	1	.8940	.1060	.1703	.1430	.9096	.0904	.1119	.2482	.9136	.0864	.0911	.7610
μ_γ	0	-.0314	.0314	.1022	.0722	-.0299	.0299	.0567	.0902	-.0281	.0280	.0358	.2402
σ_γ	.5	.4635	.0365	.0821	.1208	.4700	.0300	.0481	.1468	.4730	.0270	.0317	.3390
ρ	.5	.3851	.1149	.2459	.0708	.4074	.0926	.1371	.0916	.4088	.0912	.1014	.4904
Midpoint estimates - DGP4													
		25 obs.				100 obs.				500 obs.			
coefficient	true value	mean	bias	rMSE	size	mean	bias	rMSE	size	mean	bias	rMSE	size
μ_α	-.5	-.6737	.1737	.2263	.2162	-.6747	.1748	.1895	.6814	-.6734	.1734	.1763	1
σ_α	1	.7042	.2958	.3140	.7666	.7231	.2769	.2820	.9986	.7263	.2737	.2747	1
μ_γ	0	-.1430	.1430	.1600	.5132	-.1434	.1434	.1478	.9864	-.1430	.1430	.1439	1
σ_γ	.5	.3501	.1499	.1570	.8156	.3567	.1433	.1453	.9996	.3587	.1413	.1417	1
ρ	.5	.2073	.2927	.3496	.2866	.2178	.2822	.2954	.8948	.2164	.2836	.2861	1

reason for this is that the corresponding distribution generates samples characterised by observations mostly concentrated in the open-ended intervals which raise the discussed identification issues. In effect, in this case the means of α and γ are 1 and 1.13, respectively,²⁴ so that a proportion of the observations are located outside the thick grid determined by the two lists.

²⁴The mean of a lognormal distribution is obtained from the mean and standard deviation of the underlying normal distribution as $\exp(\mu + \sigma^2/2)$.

Presenting this example we want to make a case for the inappropriateness of this specific lists to convey information in a more extreme (but not impossible) situation, which would instead require redesign of the lists to ensure a more extensive coverage of the areas towards which the parameters, and consequently the observations, may be likely to gravitate.

8 Conclusion

MPL choice series and its “switching” variant are popular tools for eliciting preferences (mainly to risk) from subjects either in the laboratory or in the field.

In this work, we have presented an estimator for data collected via sMPL series meant to estimate multiple-parameter choice models. We have derived the estimator for the bivariate case of the Rank Dependent Utility function for choices under risk. However, there is no limit to its applicability. For example, it may be used to estimate social preference functions of the popular Fehr and Schmidt type (see Fehr and Schmidt 1999).

By a Monte Carlo simulation, we have assessed its small-sample properties, and we have established that it is eminently superior to the midpoint estimator, which has made several appearances in the literature.

Far be it from us to maintain that this is the “best” possible estimator of this sort of data, we conclude that our estimator is certainly 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.

References

- Andersen, S., Harrison, G., Hole, A., Lau, M., and Rutström, E. (2012). Non-linear mixed logit. *Theory and Decision*, 73(1):77–96.
- Andersen, S., Harrison, G., Lau, M., and Rutström, E. (2006). Elicitation using multiple price list formats. *Experimental Economics*, 9(4):383–405.
- Bosch-Domènech, A. and Silvestre, J. (2013). Measuring risk aversion with lists: a new bias. *Theory and Decision*, 75(4):465–496.
- Bruhin, A., Fehr-Duda, H., and Epper, T. (2010). Risk and rationality: Uncovering heterogeneity in probability distortion. *Econometrica*, 78(4):1375–1412.
- Conte, A., Hey, J., and Moffatt, P. (2011). Mixture models of choice under risk. *Journal of Econometrics*, 162(1):79–88.
- Conte, A. and Hey, J. D. (2018). Rehabilitating the Random Utility Model. A comment on Apestegui and Ballester (2018). Discussion Papers 18/12, Department of Economics, University of York.
- Conte, A. and Moffatt, P. (2014). The econometric modelling of social preferences. *Theory and Decision*, 76(1):119–145.
- Fehr, E. and Schmidt, K. M. (1999). A Theory of Fairness, Competition, and Cooperation. *The Quarterly Journal of Economics*, 114(3):817–868.
- Freeman, D. J., Halevy, Y., and Kneeland, T. (2019). Eliciting risk preferences using choice lists. *Quantitative Economics*, 10(1):217–237.
- Gouriéroux, C. and Monfort, A. (1996). *Simulation-based Econometric Methods*. OUP/CORE Lecture Series, Oxford University Press.
- Harrison, G. W., Lau, M. I., and Williams, M. B. (2002). Estimating individual discount rates in denmark: A field experiment. *American Economic Review*, 92(5):1606–1617.
- Holt, C. A. and Laury, S. K. (2002). Risk Aversion and Incentive Effects. *American Economic Review*, 92(5):1644–1655.
- 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.
- Loomes, G., Moffatt, P. G., and Sugden, R. (2002). A microeconomic test of alternative stochastic theories of risky choice. *Journal of Risk and Uncertainty*, 24(2):103–130.
- Loomes, G. and Sugden, R. (1995). Incorporating a stochastic element into decision theories. *European Economic Review*, 39(3):641–648. Papers and Proceedings of the Ninth Annual Congress European Economic Association.
- McFadden, D. and 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 Griliches, Z. and Intriligator, M. D., editors, *Handbook of Econometrics*, volume 2 of *Handbook of Econometrics*, chapter 24, pages 1395–1457. Elsevier.
- Morio, J., Balesdent, M., Jacquemart, D., and Vergé, C. (2014). A survey of rare event simulation methods for static input-output models. *Simulation Modelling Practice and Theory*, 49:287–304.

- 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.
- Riddel, M. (2012). Comparing risk preferences over financial and environmental lotteries. *Journal of Risk and Uncertainty*, 45(2):135–157.
- Riddel, M. and Kolstoe, S. (2013). Heterogeneity in life-duration preferences: Are risky recreationists really more risk loving? *Journal of Risk and Uncertainty*, 46(2):191–213.
- Rubinstein, R. Y. and Kroese, D. P. (2017). *Simulation and the Monte Carlo Method*. John Wiley & Sons, Inc., New York, NY, USA, 3rd edition.
- 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.
- Tanaka, T., Camerer, C. F., and Nguyen, Q. (2010). Risk and time preferences: Linking experimental and household survey data from vietnam. *American Economic Review*, 100(1):557–71.
- Train, K. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge, UK, 2nd edition.
- Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323.
- von Gaudecker, H.-M., van Soest, A., and Wengstrom, E. (2011). Heterogeneity in risky choice behavior in a broad population. *American Economic Review*, 101(2):664–94.

Appendices

SUPPLEMENTAL MATERIAL FOR ONLINE PUBLICATION

These appendices are dedicated to those who do not have familiarity with interval regression and Monte Carlo integration.

A Estimating a single-parameter preference function

Suppose we have a sample of observations from List 1 in Table 2. We want to estimate a one-parameter preference function assuming that subjects are EU maximisers with the power-CRRA utility function $u_i(z) = z^{\theta_i}$, with $\theta_i > 0$. Neglecting the apex $L = 1$ for the list, the expected utilities of the two lotteries are computed as

$$EU_i(A) = EU_i(a_1, a_2; p|\theta_i) = pu_i(a_1) + (1 - p)u_i(a_2) \quad (\text{A.1})$$

$$EU_i(B^t) = EU_i(b_1^t, b_2; q|\theta_i) = qu_i(b_1^t) + (1 - q)u_i(b_2)$$

Suppose that subject i switches to lottery B in row 6 in the series, that is $\tilde{t} = 6$. This implies that i prefers lottery A to lottery B in the first 5 rows and the other way round from row 6 on. In this case, the meaningful information is provided by the two rows in the following table.

Row	Lottery A				Lottery B			
	probability (\$400)	payoff	probability (\$100)	payoff	probability (\$?)	payoff	probability (\$50)	payoff
5	0.3	\$400	0.7	\$100	0.1	\$1,060	0.9	\$50
6	0.3	\$400	0.7	\$100	0.1	\$1,250	0.9	\$50

The values of θ_i which may have generated that choice have to satisfy conditions (2), that is

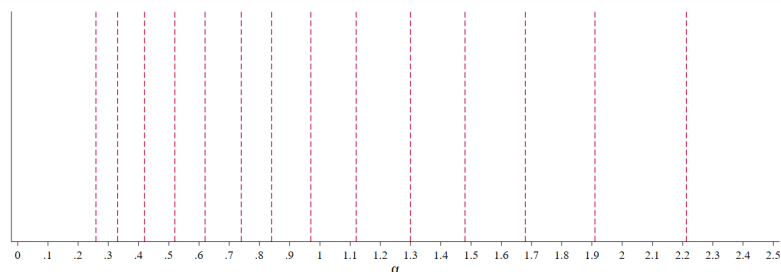
$$EU_i(B^6) > EU_i(A) > EU_i(B^5)$$

or equivalently

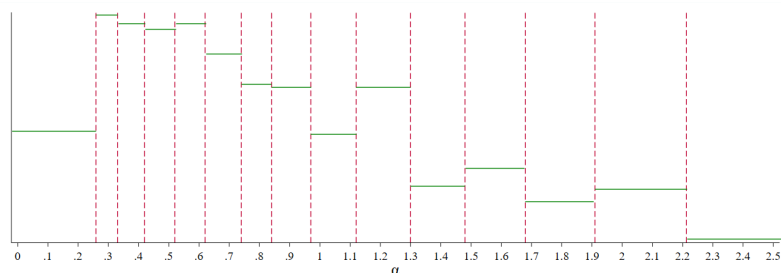
$$EU_i(1250, 50; 0.1|\theta_i) > EU_i(400, 100; 0.3|\theta_i) > EU_i(1060, 50; 0.1|\theta_i)$$

The solution to these conditions and the conditions implied by the other possible switch points produces the following figure, where the vertical bars indicate the upper and lower bounds of θ_i for any possible choice of the switch point. We note that switching in row 1 is determined by all

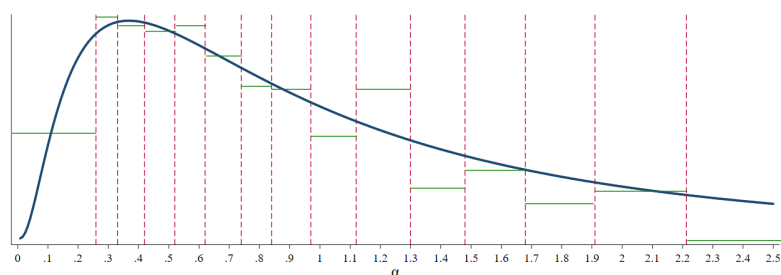
the values in the right open bound. Switching in row 2 is implied by the values of θ between the last vertical bar on the right and the penultimate, and so on.



Suppose that we have some observations collected from a sample of subjects faced with List 1. We can imagine to place such observations over the previous plot as in the following one, where the areas underneath the horizontal and within the vertical bars represent the proportion of observations in those areas.



Estimating θ from a sample of observations means finding an estimator of the probability of choosing each switch point. In other words, the estimated distribution has to be such that, if a random sample is drawn from it, this random sample is able to “mimic” the characteristics of the data or such that the available sample would not look unusual if drawn from the estimated distribution. An example is given by the density plotted in the following figure. The estimation procedure is based on finding the probability mass within the bounds of θ for each switch point which is estimated by the area underneath the density function delimited by such bounds.



A.1 Distributional assumptions and likelihood contribution

Technically, the variable of interest θ that we want to estimate is intrinsically continuous, but it is as if we observed it through a filter, because it has been discretised into intervals of values that contain the actual values. The intervals are represented by the vertical bars in the plots above.

Once known that the value of θ for subjects i is somewhere in the interval $(\underline{\theta}_i; \bar{\theta}_i)$, the likelihood contribution from this individual is simply

$$\Pr(\underline{\theta}_i \leq \theta \leq \bar{\theta}_i)$$

which becomes $\Pr(\theta \leq \bar{\theta}_i)$ and $\Pr(\underline{\theta}_i \leq \theta)$ in the cases in which the lower limit ($\underline{\theta}_i$) and the upper limit ($\bar{\theta}_i$) of the intervals are $-\infty$ and ∞ , respectively, that is in the open-ended cases.

To compute this probability, we need to make an assumption about the distribution of θ . It is important to respect the natural domain of the parameter to estimate. In our case, we have to choose a distribution taking values from 0 to ∞ , since in the chosen utility function θ can only take positive values.¹

Let $h(\theta)$ and $H(\theta)$ denote θ_i 's probability density and distribution functions evaluated at θ , respectively. Essentially, that probability corresponds to the area underneath the density function within $\underline{\theta}_i$ and $\bar{\theta}_i$, that is

$$\begin{aligned} \Pr(\underline{\theta}_i \leq \theta \leq \bar{\theta}_i) &= H(\bar{\theta}_i) - H(\underline{\theta}_i) \\ &= \int_{\underline{\theta}_i}^{\bar{\theta}_i} h(\theta) d\theta = \int_0^{\infty} \mathbf{I}[\underline{\theta}_i \leq \theta \leq \bar{\theta}_i] h(\theta) d\theta \end{aligned}$$

This is the individual contribution that is passed to the likelihood function to maximise. We want to draw the reader's attention on the equivalent ways in which this probability can be written, and especially on the differences between the bottom two, because it is going to be useful in what follows.

When the distribution of the parameter of interest can be chosen to be normal, this problem solves in the standard "interval regression model", which is available in the most popular statistical packages. If, instead, the distribution of the parameter can be assumed to be lognormal, as in our example, it is sufficient to take the natural logarithm of the upper and lower bounds of each choice and apply the interval regression on the transformed bounds. This is possible because of

¹Often, the suitable distribution can be suggested by the data itself, which may show some peculiarities (such as multi-modality, a mass at a particular value, etc.) that the distribution to estimate has to be able to replicate.

the way these two distributions are related.

A.2 Using Monte Carlo integration

Suppose that $h(\theta)$ is not normal or lognormal. The $\Pr(\underline{\theta}_i \leq \theta \leq \bar{\theta}_i)$ can be easily programmed by the user, provided that the probability distribution function $H(\theta)$ is available in the statistical package. Otherwise, one can still resort to using Monte Carlo integration (see Train 2009, among others).

Let us focus on this formulation of the likelihood contribution

$$\Pr(\underline{\theta}_i \leq \theta \leq \bar{\theta}_i) = \int_0^\infty \mathbf{I}[\underline{\theta} \leq \theta \leq \bar{\theta}_i] h(\theta) d\theta$$

Suppose we can draw an *arbitrarily* large number of θ 's, denoted as $\hat{\theta}_r$ with $r = 1, \dots, R$ from its distribution $h(\theta)$. For every such value, we compute the indicator function $\mathbf{I}[\underline{\theta} \leq \theta \leq \bar{\theta}_i]$, obtaining a sequence of 0's and 1's. Finally, we calculate the proportion of 1's, which, for $R \rightarrow \infty$, approximates or better estimates the integral above

$$\Pr(\underline{\theta}_i \leq \theta \leq \bar{\theta}_i) \approx \frac{\sum_{r=1}^R \mathbf{I}[\underline{\theta} \leq \hat{\theta}_r \leq \bar{\theta}_i]}{R}$$

Therefore, the “crude frequency estimator” of the integral is

$$\Pr(\widehat{\underline{\theta}_i \leq \theta \leq \bar{\theta}_i}) = \frac{\sum_{r=1}^R \mathbf{I}[\underline{\theta} \leq \hat{\theta}_r \leq \bar{\theta}_i]}{R}, \quad i = 1, \dots, n$$

This estimator has the same drawbacks in the single-parameter case discussed in Section 6 for this same estimator in the bivariate case. Therefore, it requires that another technique for drawing from densities is used (see Stern 1997 for further details).

A.2.1 Extension to a two-parameter function

As extensively discussed in Section 6, extending the interval regression model to the bivariate case at hand is not so straightforward. It is actually impossible for the following reason. Consider equation (4) and consider writing that probability using the lower and upper bound of the parameters as in the univariate case, that is

$$\begin{aligned}
\Pr(\tilde{t}^1 = h \& \tilde{t}^2 = k) &= \int_0^\infty \int_0^\infty \mathbf{1}[C_i^1(h) \& C_i^2(k)] f(\alpha, \gamma) d\alpha d\gamma \\
&= \int_{\underline{\alpha}_i}^{\bar{\alpha}_i} \int_{\underline{\gamma}_i}^{\bar{\gamma}_i} f(\alpha, \gamma) d\alpha d\gamma
\end{aligned}$$

This would be a simple extension of the univariate case just discussed, but would be incorrect. In effect, the bottom integral calculates the probability mass within the red lines in Figure 2, overestimating that within the (α, γ) which is strictly smaller. Using a bivariate interval regression, which is fine with rectangular probabilities, in this case would considerably bias the estimates.

B The midpoint estimator

Given a choice of switch points by subject i , let $\check{\alpha}_i$ and $\check{\gamma}_i$ the abscissa and ordinate of the midpoint corresponding to that choice as in Figure 1, bottom-right panel, calculated as described in Section 5. For each subject in the sample, we obtain a pair of such values. In our case, given our distributional assumptions concerning α and γ , we can write the model as

$$\begin{cases} \ln(\check{\alpha}_i) = \mu_\alpha + \sigma_\alpha \epsilon \\ \ln(\check{\gamma}_i) = \mu_\gamma + \sigma_\gamma v \end{cases}, \quad \begin{pmatrix} \epsilon \\ v \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$$

This model can be estimated by Maximum Likelihood, where the likelihood contribution of subject i is

$$h(\ln(\check{\alpha}_i), \ln(\check{\gamma}_i)) = \frac{1}{2\pi\sigma_\alpha\sigma_\gamma\sqrt{1-\rho^2}} \exp\left[-\frac{z}{2(1-\rho^2)}\right]$$

where

$$z \equiv \frac{(\check{\alpha}_i - \mu_\alpha)^2}{\sigma_\alpha^2} - \frac{2\rho(\check{\alpha}_i - \mu_\alpha)(\check{\gamma}_i - \mu_\gamma)}{\sigma_\alpha\sigma_\gamma} + \frac{(\check{\gamma}_i - \mu_\gamma)^2}{\sigma_\gamma^2}$$

The estimator using such information is what we named “midpoint estimator”. When the distribution of the parameters of interest or their logarithm is bivariate normal, the estimator is the bivariate extension of the linear regression model estimator (known as SURE).