



HAL
open science

Eliciting Multiple Prior Beliefs

Brian Hill, Mohammed Abdellaoui, Philippe Colo

► **To cite this version:**

Brian Hill, Mohammed Abdellaoui, Philippe Colo. Eliciting Multiple Prior Beliefs. 2021. hal-03503996

HAL Id: hal-03503996

<https://hal.science/hal-03503996>

Preprint submitted on 28 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Eliciting Multiple Prior Beliefs*

Mohammed Abdellaoui
HEC Paris & CNRS[†]

Philippe Colo
PSE, Université Paris I
IESEG[‡]

Brian Hill
HEC Paris & CNRS[§]

3rd June 2021

Abstract

Despite the increasing importance of multiple priors in various domains of economics and the significant theoretical advances concerning them, choice-based incentive-compatible multiple-prior elicitation largely remains an open problem. This paper develops a solution, comprising a preference-based identification of a subject's probability interval for an event, and two procedures for eliciting it. The method does not rely on specific assumptions about subjects' ambiguity attitudes or probabilistic sophistication. To demonstrate its feasibility, we implement it in two incentivized experiments to elicit the multiple-prior equivalent of subjects' cumulative distribution functions over continuous-valued sources of uncertainty. We find a predominance of non-degenerate probability intervals among subjects for all explored sources, with intervals being wider for less familiar sources. Finally, we use our method to undertake the first elicitation of the mixture coefficient in the Hurwicz α -maxmin EU model that fully controls for beliefs.

Key words: Multiple Priors, Belief Measurement, α -maxmin EU, Imprecise Probability.

JEL Codes: D81

*Acknowledgements and thanks to be added. This research is supported by grants of the French National Research Agency (ANR), "Investissements d'Avenir" (LabEx Ecodec/ANR-11-LABX-0047) (ANR) and the ANR project DUSUCA (ANR-14-CE29-0003-01).

[†]1 rue de la Libération, 78351 Jouy-en-Josas, France. E-mail: abdellaoui@hec.fr.

[‡]48 Boulevard Jourdan, 75014, Paris, France. E-mail: colo.philippe@gmail.com.

[§]1 rue de la Libération, 78351 Jouy-en-Josas, France. E-mail: hill@hec.fr.

20 Contents

21	1 Introduction	4
22	2 Theoretical Background	6
23	2.1 Preliminaries	6
24	2.2 Upper and lower probabilities and CDFs	7
25	2.3 Decision model	8
26	2.4 Theory: Matching Probability Intervals	9
27	2.5 Implementation 1: 2D Choice lists	11
28	2.6 Implementation 2: Binary-choice procedure	12
29	3 Experimental Methods	14
30	3.1 Subjects	14
31	3.2 Sources of uncertainty	15
32	3.3 Choice Tasks	16
33	3.4 Elicitation procedures	17
34	3.5 Incentivizing subjects	18
35	4 Results	19
36	4.1 Descriptive Statistics and Performance	19
37	4.2 Bayesian analysis	20
38	4.3 Imprecision	22
39	4.4 Imprecision and familiarity	25
40	4.5 Matching probabilities and the α -maxmin EU mixture coefficient	26
41	5 Discussion	26
42	6 Conclusion	31
43	A Theoretical Appendix	33
44	A.1 Proofs	33
45	A.2 Binary-choice procedure	34
46	A.2.1 Introduction and setup	34
47	A.2.2 Presentation of main steps	35
48	A.2.3 Convergence	36
49	A.3 Robustness of the elicitation method	42
50	A.3.1 Matching Probability Intervals	42
51	A.3.2 2D Choice List	44
52	A.3.3 Binary-choice procedure	45

53	B Supplementary Statistics	46
54	B.1 Descriptive Statistics	46
55	B.1.1 Imprecision	46
56	B.2 Bayesian estimation	54
57	B.2.1 Statistical approach	54
58	B.2.2 Analysis	55
59	B.3 Matching Probability data and analysis of α	60
60	B.4 Elicitation-free check of $\alpha > \frac{1}{2}$	60
61	C Experimental design and displays	63
62	C.1 Probability interval elicitation: displays	63
63	C.2 MP elicitation: displays	66
64	D Technical Appendix: Lemmas for the proof of Proposition 4	68

1. Introduction

The standard Bayesian model of decision under uncertainty in economics stipulates that a decision maker’s beliefs are fully captured by a single probability measure over the states of the world (Savage, 1954; Anscombe and Aumann, 1963). However, in the face of contrary empirical evidence, starting with Ellsberg (1961)’s famous examples, more general theories have weakened the standard assumption of probabilistic beliefs. One of the most popular models involves as a ‘belief component’ a set of priors over the states of the world (Gilboa and Schmeidler, 1989). Multiple prior models have found a growing number of applications in macroeconomics (Ilut and Schneider, 2014), finance (Garlappi et al., 2007; Epstein and Schneider, 2010), mechanism design (Bose and Renou, 2014), econometrics (Manski, 2003, 2013), health economics (Giustinelli et al., 2021), but also beyond economics, in climate science (Kriegler et al., 2009), risk analysis (Cox, 2012) and uncertainty communication (Dieckmann et al., 2010), including by central banks (Carney et al., 2019). Despite obvious theoretical appeal, empirical applications of multiple prior models still have to operate in the absence of appropriate choice-based procedures for eliciting their ‘belief component’. To date, almost all attempts to operationalize multiple prior elicitation have focused on subjects’ *stated* probability intervals for individual events (Giustinelli et al., 2021; Kriegler et al., 2009), and hence involve procedures that are neither choice based nor incentive compatible. This paper proposes a choice-based incentive-compatible elicitation method for probability intervals, and implements it in two laboratory experiments.

Our proposal is inspired by the matching probability (MP) method for determining the subjective probability of an uncertain event (Borel, 1939; Anscombe and Aumann, 1963). Under Subjective Expected Utility (SEU), the subjective probability of a target event E coincides with its MP, which can be inferred from preferences between a bet on E and bets on events generated by extraneous random devices with known probability, e.g. the colour of a randomly drawn ball from an urn of known composition. Specifically, for urns containing only red or blue balls, the MP is given by the proportion r of red balls such that the subject is indifferent between the gamble that pays out a monetary prize z if E occurs, and nothing otherwise, and the gamble on the urn with proportion r of red balls that pays z if the next ball drawn from the urn is red (Abdellaoui et al., 2005; Dimmock et al., 2015).

Our insight for eliciting probability intervals is to use extraneous random devices with *interval-valued* rather than precise probabilities. To illustrate, consider an urn containing only red and blue balls, where all that is known is that at least proportion r of the balls in the urn are red, and at least proportion b are blue (with $r + b \leq 1$). Here, the probabilities of getting red or blue on the next draw from the urn are summarized by the intervals $[r, 1 - b]$ and $[b, 1 - r]$, respectively. Under the popular Hurwicz α -maxmin expected utility model, a single indifference between the gamble on the target event E and bets on such urns does not suffice to identify the subject’s probability interval for E . However, we show that the latter can be identified from a *pair* of correctly-chosen indifferences involving bets concerning E and such ‘interval-valued’

urns. As explained in detail in Section 2.4, we thus propose a preference-based association of an ‘interval-valued’ urn to each event, which coincides with the subject’s probability interval for the event under Hurwicz α -maxmin expected utility. This *matching probability interval* notion resolves the problem of choice-based incentive-compatible probability-interval elicitation in theory.

Our elicitation approach is theoretically robust, as can be seen in the weakness of the assumptions underlying it. On the one hand, it operates under Hurwicz α -maxmin expected utility, and hence under a range of ambiguity attitudes. In particular, this model is consistent with recent experimental evidence suggesting that people may exhibit ambiguity aversion in some choices and be ambiguity seeking in others (Kocher et al., 2018). On the other hand, our approach does not require that subjects’ preferences—or sets of priors—are generated by precise probabilistic beliefs, i.e. that they are probabilistically sophisticated (Machina and Schmeidler, 1992; Chew and Sagi, 2006). Multiple priors clearly come to the fore in situations where assumptions of this sort are unwarranted. Elicitation under Hurwicz α -maxmin expected utility in the absence of probabilistic sophistication faces well-known theoretical difficulties with the identification of this model (e.g. Ghirardato et al., 2004; Eichberger et al., 2011), and one contribution of our approach is to provide a new resolution of them (Section 5). Hill (2021) develops the axiomatic underpinnings of an extension of the approach taken here, in a more general setup.

We then operationalise elicitation of matching probability intervals in a laboratory setting. Since probability intervals are two-dimensional, standard techniques for eliciting MPs, which are one-dimensional, cannot be used. We develop two methods for eliciting matching probability intervals: a two-dimensional extension of well-known (one-dimensional) choice lists, and an adaptive binary-choice procedure, which can be thought of as an interval analogue of the bisection or staircase method for eliciting matching probabilities. We implement these methods in tandem, following the hybrid elicitation approach adopted by Abdellaoui et al. (2021). Many elicitation applications in economics and beyond require subjects’ probability distributions or cumulative distribution functions (CDFs) over a continuous variable of interest (e.g. US inflation in 2023, Eurozone GDP in 2022, average global temperature in 2030). Motivated by this observation, we implement our methods on two pairs of sources of uncertainty of the latter sort, to elicit the interval-valued CDFs generated by subjects’ multiple priors. Interval-valued CDFs are commonly used in applications to go beyond the assumption of precise subjective probabilities (Karanki et al., 2009); our elicitation of CDFs provides a test of our approach, showing that it can operate in such contexts.

Our central findings attest to the feasibility of the approach. Our method yields generally consistent results, eliciting, for the vast majority of subjects, non-degenerate interval-valued CDFs. Our elicitation suggests that imprecise beliefs—i.e. non-singleton intervals for some events—are widespread, with only a handful of subjects having fully precise probabilities for all elicited events. This finding, which is consistent with elicitation using stated probability intervals (Giustinelli et al., 2021), attests to the relevance of multiple-prior belief elicitation.

142 Moreover, by eliciting subjects' beliefs for two similar sources which intuitively differ in famil-
143 iarity or predictability (e.g. the temperature in Paris and in Sydney for subjects in Paris),
144 our elicitations provide insight into the relationship between intuitive familiarity or predictabil-
145 ity and probability intervals. Specifically, we observe that the width of subjective probability
146 intervals is typically larger for intuitively less familiar or less predictable sources. Again, the
147 reasonableness of this correlation corroborates the solidity of our method.

148 Finally, we connect our elicited beliefs with the Hurwicz α -maxmin EU model, and perform
149 what to our knowledge is the first elicitation of the mixture coefficient α in that model that fully
150 controls for beliefs without making particular assumptions about their form, such as probabilistic
151 sophistication.

152 The paper is structured as follows. Section 2 sets out the theoretical background and
153 presents the central planks of our methods (the 'matching probability interval' notion, the two-
154 dimensional choice lists and the binary-choice procedure), with the relevant theoretical results.
155 Section 3 sets out our experimental implementations, in the form of two studies. Section 4
156 contains our results and supporting analyses, whereas in Section 5 we discuss connected is-
157 sues, related literature and future directions. Proofs, data analyses and experimental details are
158 contained in the Appendices.

159 2. Theoretical Background

160 In this section, we first set out the general setup, the objects of elicitation and the underlying
161 decision model (Sections 2.1–2.3). Then we present in turn the elements of our method. First,
162 we propose an analogue of MPs for probability intervals, and show that they are sufficient to
163 yield the subject's probability interval for an event, in theory (Section 2.4). Then we turn to im-
164 plementation, presenting an extended notion of choice list for probability intervals (Section 2.5)
165 and a binary-choice procedure, reminiscent of the bisection procedure for matching probabilities
166 (Section 2.6).

167 2.1. Preliminaries

168 We consider decision making situations where the objects of choice are two-outcome prospects
169 that pay a fixed monetary outcome z if an event occurs, and nothing otherwise. Prospects
170 with general winning event E and winning amount z are denoted $(z, E, 0)$ and called *bets*. The
171 *complementary* bet, which pays out when the event E does not occur, is denoted $(0, E, z)$.

172 As mentioned previously, we use extraneous interval-valued random devices realised by urns
173 containing red and blue balls with partial information about the composition. For instance,
174 consider the urn where subjects are only told that at least a proportion r of its balls are red,
175 at least a proportion b are blue (with $r + b \leq 1$) but receive no information about the colour
176 composition of the remaining balls (except that each is either red or blue). For such an urn,
177 the information only allows assignment of the interval $[r, 1 - b]$ for the probability of the next

178 ball drawn from the urn being red; similarly, there is the interval $[b, 1 - r]$ for the next ball
 179 being blue. Using these probability intervals for parametrization, we denote the urn with at
 180 least proportion r of red balls and at least proportion b of blue balls by $[r, 1 - b]$. We denote the
 181 set of such *interval-valued urns* by \mathcal{I} .¹

182 Each urn $[r, 1 - b]$ in \mathcal{I} can be related to two (sorts of) prospects. One is the prospect which
 183 pays z if the next ball drawn from the urn is red, and nothing otherwise. For such a prospect,
 184 the probability of winning is characterized by the interval $[r, 1 - b]$; we denote this prospect by
 185 $(z, [r, 1 - b], 0)$. The other prospect involves the complementary bet on this urn—that is, the
 186 bet on the next ball drawn from it being blue. We denote this prospect by $(0, [r, 1 - b], z)$. Note
 187 that the probability of *losing* here is characterised by the interval $[r, 1 - b]$, so the probability of
 188 winning is given by $[b, 1 - r]$; this prospect is thus essentially equivalent to $(z, [b, 1 - r], 0)$. Since
 189 these prospects all involve objectively given information about the probability of winning, albeit
 190 in interval rather than precise probability form, we call them *interval lotteries* (IL).² Standard
 191 lotteries correspond to the special case where the composition of the urn is fully known—i.e.
 192 $r = 1 - b$. So, for instance, the *matching probability* (MP) of an event E can be defined in this
 193 setup as the r such that $(z, [r, r], 0) \sim (z, E, 0)$.

194 The set \mathcal{I} of interval-valued urns can be visually represented by the black-edged triangle in
 195 Figure 1. Each point (x, y) represents the urn $[x, y]$ —i.e. with at least proportion x of red balls
 196 and at least proportion $1 - y$ of blue ones. As such, it represents two interval lotteries: the bet
 197 on red, $(z, [x, y], 0)$, where all that is known is that the winning probability is in the interval
 198 $[x, y]$, and the bet on blue $(0, [x, y], z)$, where all that is known is that the losing probability is
 199 in this interval. Standard lotteries and urns with fully known composition correspond to the
 200 points on the diagonal $(x = y)$.

201 2.2. Upper and lower probabilities and CDFs

202 The sources of uncertainty considered here are real-valued variables, e.g. the daily minimum
 203 temperature in Paris between November and March. In the precise probability case, elicitation
 204 aims at revealing the subjective probability over the variable, which can be represented as a
 205 subjective cumulative distribution function (CDF). One common way of doing so is by eliciting
 206 subjective probabilities of events corresponding to the variable lying below certain fixed values.
 207 For a variable taking values in a real interval T , the events considered are of the form $E_t =$
 208 $\{t' \in T : t' \leq t\}$. For future reference, we call these *cumulative events*. We now set out the aim
 209 of the corresponding exercise for multiple priors.

210 Multiple prior belief representations involve a convex, closed set \mathcal{C} of probability mea-

¹Formally: $\mathcal{I} = \{[x, y] : (x, y) \in \mathbb{R}^2, 0 \leq x \leq y \leq 1\}$.

²Our notion of interval lottery is distinct from that used by Gul and Pesendorfer (2014). They use ‘interval lottery’ to denote (precise) probability measures over the set of intervals of (monetary) prizes; here, ‘interval lottery’ denotes assignments of probability intervals to (fully determined, precise) outcomes. In particular, the interval lotteries $(z, [r, 1 - b], 0)$ used here clearly do not belong to the concept used by Gul and Pesendorfer (zero probability is assigned to each outcome in the interior of the interval $[0, z]$).

211 sures: measures over the values of variable of interest, in our case.³ For each event E_t , the
 212 set of priors generates a *probability interval* $\{p(E_t) : p \in \mathcal{C}\} = [\underline{p}(E_t), \bar{p}(E_t)]$, where $\underline{p}(E_t) =$
 213 $\min \{p(E_t) : p \in \mathcal{C}\}$ and $\bar{p}(E_t) = \max \{p(E_t) : p \in \mathcal{C}\}$ are the *lower* and *upper probabilities* for
 214 E_t respectively. As is well-known, a set of priors contains more information than the collection of
 215 upper and lower probabilities for all events generated from it, but the latter (or sometimes less) is
 216 often sufficient for applications, and sometimes preferable insofar as it is easier to communicate.

217 For continuous-valued variables, CDFs are often used. Recall that for a probability measure
 218 $p \in \Delta(T)$, the CDF is defined as $F_p(t) = p(\{t' \in T : t' \leq t\}) = p(E_t)$; when the probability
 219 measure is a subjective probability, this is the corresponding subjective CDF. In this context,
 220 a set of priors \mathcal{C} generates the *interval-valued CDF* $F_{\mathcal{C}}(t) = \{p(E_t) : p \in \mathcal{C}\}$, which takes the
 221 probability interval corresponding to E_t as value, for each t . This can be visually represented
 222 in terms of two (real-valued) functions: the *lower CDF*, $\underline{F}_{\mathcal{C}}(t) = \min \{p(E_t) : p \in \mathcal{C}\} = \underline{p}(E_t)$,
 223 and the *upper CDF*, $\overline{F}_{\mathcal{C}}(t) = \max \{p(E_t) : p \in \mathcal{C}\} = \bar{p}(E_t)$. Although, like upper and lower
 224 probabilities, they involve an information loss as compared to sets of priors, these are widely
 225 used for representing, communicating and studying sets of priors over continuous variables, where
 226 they often go under the name of distribution bands or p-boxes (Berger et al., 2000; Karanki et al.,
 227 2009). In the implementation of our elicitation procedure conducted here, our aim is to elicit
 228 subjective upper and lower CDFs for the variables considered.

229 2.3. Decision model

230 For the purposes of presentation, we will focus on one of the most popular and general models
 231 of decision under uncertainty involving sets of priors, the Hurwicz α -maxmin EU model. (In
 232 Section 5 and Appendix A.3, we discuss how our proposals extend to generalisations.) Under
 233 the α -maxmin model, a bet $(z, E, 0)$ is evaluated according to:

$$\alpha \underline{p}(E).u(z) + (1 - \alpha) \bar{p}(E).u(z) \tag{1}$$

234 where $\bar{p}(E)$ and $\underline{p}(E)$ are the upper and lower probabilities of E generated by the subjects' set of
 235 priors, as defined above, and u is a utility function normalized so that $u(0) = 0$. The coefficient
 236 α is often taken to reflect ambiguity attitude in this model, with $\alpha > \frac{1}{2}$ associated with typical
 237 ambiguity aversion and $\alpha < \frac{1}{2}$ with typical ambiguity seeking. For illustration, the standard
 238 behavior in the Ellsberg two-urn example can be accommodated by $\alpha > \frac{1}{2}$ but not by $\alpha < \frac{1}{2}$.
 239 This model coincides with the Gilboa-Schmeidler maxmin-EU model when $\alpha = 1$; whenever
 240 $\alpha \neq 1$, the model does not satisfy the Gilboa-Schmeidler uncertainty aversion axiom—which
 241 can be thought of as characterizing universal ambiguity aversion. Hence, it can accommodate
 242 ambiguity seeking behavior in certain choices (even for $1 > \alpha > \frac{1}{2}$). Since typical findings
 243 suggest some ambiguity seeking behavior, but not in situations that give reason to believe that
 244 $\alpha < \frac{1}{2}$, we take $\alpha > \frac{1}{2}$ to be typical and assume that preferences are represented according to (1)
 245 with $\alpha > \frac{1}{2}$ in the sequel (except where specified). As discussed in Section A.3, the full strength

³Technically, $\mathcal{C} \subseteq \Delta(T)$, the set of probability measures over T .

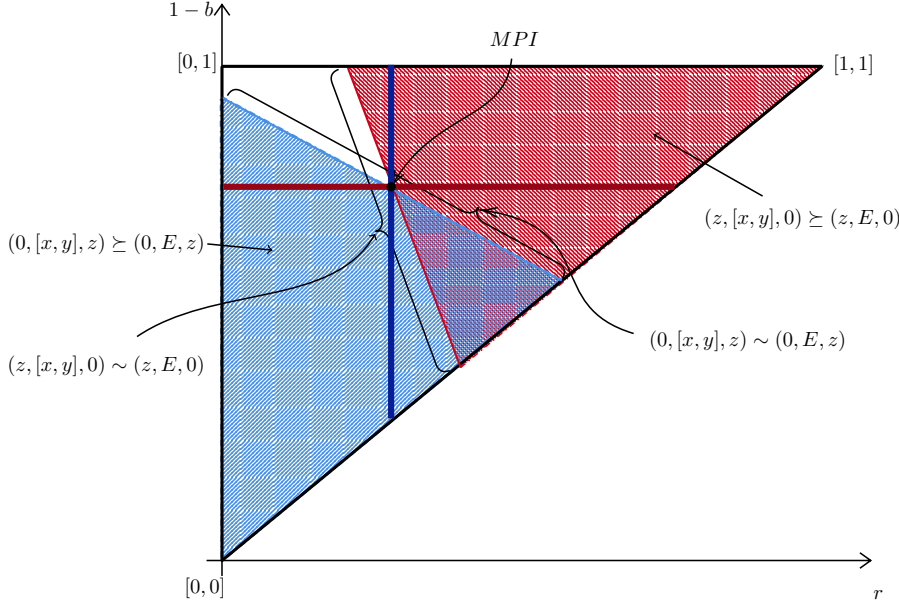


Figure 1: Matching Probability Interval in space \mathcal{I} of interval-valued urns, for an event E .

246 of this assumption is not required for the central elements of the elicitation method. Our aim is
 247 to elicit $\underline{p}(E)$ and $\bar{p}(E)$ for selected events.

248 We also assume the same representation for interval lotteries: i.e. preferences concerning
 249 them are represented by an evaluation of $(z, [r, 1 - b], 0)$ by:

$$\alpha r u(z) + (1 - \alpha)(1 - b)u(z) \quad (2)$$

250 2.4. Theory: Matching Probability Intervals

Our approach is based on the following notion. The *matching probability interval (MPI)* of an event E is an $[r, 1 - b] \in \mathcal{I}$ such that:

$$(z, [r, 1 - b], 0) \sim (z, E, 0) \quad (3)$$

$$(0, [r, 1 - b], z) \sim (0, E, z) \quad (4)$$

251 Plugging these indifferences into (1) and (2) yields the following equations:

$$\begin{aligned} \alpha r + (1 - \alpha)(1 - b) &= \alpha \underline{p}(E) + (1 - \alpha)\bar{p}(E), \\ \alpha(1 - (1 - b)) + (1 - \alpha)(1 - r) &= \alpha(1 - \bar{p}(E)) + (1 - \alpha)(1 - \underline{p}(E)). \end{aligned} \quad (5)$$

252 Clearly, these equations are satisfied by $r = \underline{p}(E)$, $1 - b = \bar{p}(E)$. Moreover, whenever there
 253 is a unique pair $r, 1 - b$ satisfying them, then there is a unique matching probability interval,
 254 which indicates precisely the subjective probability interval for E : i.e. $[\underline{p}(E), \bar{p}(E)] = [r, 1 - b]$.
 255 Under the α -maxmin EU model with $\alpha \neq \frac{1}{2}$, the MPI is unique (Proposition 3, Appendix A.1).
 256 So to elicit the subjects' probability interval for the event E , it suffices to find the MPI of E .

257 The MPI can be illustrated in Figure 1. The red hatched area represents the upper contour
 258 set of the bet $(z, E, 0)$ in the space of interval lotteries corresponding to bets on red: that is, the

259 set of (x, y) such that $(z, [x, y], 0) \succeq (z, E, 0)$. The blue hatched area is the upper contour set of
 260 the complementary bet $(0, E, z)$ in the space of complementary ILs (corresponding to bets on
 261 blue): that, it is the set of (x, y) such that $(0, [x, y], z) \succeq (0, E, z)$. The boundaries of these sets
 262 (the diagonal red and blue lines respectively) represent the indifference curves of $(z, E, 0)$ (resp.
 263 $(0, E, z)$), in the space of ‘red’ (resp. ‘blue’) ILs. The matching probability interval corresponds
 264 to the black point at the intersection of these two lines.

265 This Figure also brings out the contribution of ILs as compared to standard lotteries and the
 266 long-standing identification problem for the α -maxmin EU model (Section 5). The MP of the
 267 bet $(z, E, 0)$ is given by the point where the red indifference curve meets the diagonal; clearly
 268 eliciting it is insufficient to pin down the subject’s probability interval for E . Similarly, the MP
 269 of the complementary bet $(0, E, z)$ is given by the point where the blue indifference curve meets
 270 the diagonal. Eliciting both of these MPs is sufficient to pin down the subject’s probability
 271 interval (as the intersection of the indifference curves) *only if the slope of the indifference curves*
 272 *is known: but this is determined by the mixture coefficient α in (1), which also needs to be*
 273 *elicited*. The use of ILs, and the notion of MPI built upon it, allows elicitation of the subjective
 274 probability interval without requiring elicitation of the mixture coefficient α . Indeed, we shall
 275 use our probability interval elicitation in tandem with MPs to estimate subjects’ α (Section 4.5).

276 The notion of MPI resolves the challenge of incentive-compatible probability-interval elici-
 277 tation *in theory*. Under SEU, eliciting preferences between the bet on E and each of a range
 278 of lotteries suffices to find the MP, which corresponds to the subject’s subjective probability.
 279 Similarly here, obtaining the subject’s preferences between each pair consisting of a bet (for or
 280 against E) and an IL provides the MPI—and hence the subject’s probability interval for E —as
 281 the point satisfying (3) and (4). It is well known that there are fully incentive-compatible mech-
 282 anisms for eliciting such preferences. For instance: the subject states her preference between
 283 each pair consisting of a bet (for or against E) and an IL; a random bet (for or against E)
 284 and IL are then chosen and she is remunerated according to the prospect which she stated as
 285 more preferred between the two. By the standard argument, it is in the subject’s best interest
 286 to report her preferences truthfully, for if not there is a chance of receiving her less preferred
 287 prospect in the choice that is ‘played for real’. These elicited preferences provide, *inter alia*, the
 288 MPI and hence the subject’s probability interval for E .

289 Of course, implementation typically requires a method involving fewer preference questions.
 290 This is especially challenging for probability intervals, since the target is a point in a two-
 291 dimensional space, whereas elicitation of the precise probability of an event only needs to search
 292 a one-dimensional space. To operationalise our approach, we develop two more parsimonious
 293 methods for eliciting MPIs, whilst making no claim to have exhausted all possibilities.

294 **2.5. Implementation 1: 2D Choice lists**

Consider any MPI $[r, 1 - b]$ of an event E , so that the indifference (3) is satisfied. Since $\alpha > 0$ in the representation (1), it follows that:

$$\begin{aligned} (z, [q, 1 - b], 0) &\succ (z, E, 0) && \text{for all } q > r \\ (z, [q, 1 - b], 0) &\prec (z, E, 0) && \text{for all } q < r \end{aligned} \tag{6}$$

295 On Figure 1, this determines the preferences on the ‘red’ ILs corresponding to the bold red
296 (horizontal) line. To the left of the MPI, the bet on E is preferred to the IL corresponding to
297 the bet on red from the urn $[q, 1 - b]$ (i.e. with probability $[q, 1 - b]$ of winning); to the right of
298 the MPI, the IL is preferred to the bet; and at the MPI, the two are indifferent.

Similar reasoning applies to complementary bets. By the indifference (4) and the fact that $\alpha > 0$, for an MPI $[r, 1 - b]$ of E , it follows that:

$$\begin{aligned} (0, [r, q], z) &\prec (0, E, z) && \text{for all } q > 1 - b \\ (0, [r, q], z) &\succ (0, E, z) && \text{for all } q < 1 - b \end{aligned} \tag{7}$$

299 On Figure 1, this determines the preferences on the (complementary) ‘blue’ ILs corresponding
300 to the bold blue (vertical) line. Above the MPI, the bet against E is preferred to the IL
301 corresponding to the bet on blue from urn $[r, q]$ (i.e. with probability $[r, q]$ of losing); below the
302 MPI, the IL is preferred to the bet; and at the MPI, the two are indifferent.

303 It follows that the only points supporting the specified preference patterns on the corre-
304 sponding horizontal and vertical lines are MPIs.

Proposition 1. *For any event E , let $[r, 1 - b] \in \mathcal{I}$ be such that:*

$$\begin{aligned} (z, [q, 1 - b], 0) &\succ (z, E, 0) && \text{for all } q > r \\ (z, [q, 1 - b], 0) &\prec (z, E, 0) && \text{for all } q < r \\ (0, [r, q], z) &\prec (0, E, z) && \text{for all } q > 1 - b \\ (0, [r, q], z) &\succ (0, E, z) && \text{for all } q < 1 - b \end{aligned}$$

305 *Then $[r, 1 - b]$ is a matching probability interval of E .*

306 The red (horizontal) and blue (vertical) bold lines can thus be thought of as a pair of choice
307 lists, and the MPI is the switching point on each of them. We henceforth refer to the combination
308 of the two as a *2D choice list*. Inspired by this observation, consider the following mechanism
309 for eliciting a subject’s MPI for an event E . A subject reports an interval-valued urn $[r, 1 - b]$
310 for E . She is then remunerated as follows. First, an urn $[x, y]$ is chosen at random from the 2D
311 choice list.⁴ Then she ‘receives’ or ‘plays’ a bet or IL according to the following scheme:

- 312 • if $y = 1 - b$, $x < r$, then she gets $(z, E, 0)$ (i.e. she ‘plays’ the bet on E)

⁴I.e. the interval is chosen at random from $\{[x, y] : (x, y) \in \mathcal{I}, y = 1 - b\} \cup \{[x, y] : (x, y) \in \mathcal{I}, x = r\}$, the union of the horizontal and vertical lines going through $(r, 1 - b)$ in Figure 1.

- 313 • if $y = 1 - b$, $x \geq r$, then she gets $(z, [x, y], 0)$ (i.e. she ‘plays’ this IL)
- 314 • if $x = r$, $y < 1 - b$, then she gets $(0, [x, y], z)$ (i.e. she ‘plays’ this IL)
- 315 • if $x = r$, $y \geq 1 - b$, then she gets $(0, E, z)$ (i.e. she ‘plays’ the bet against E)

316 It follows from the previous Proposition that this mechanism is incentive compatible in the
 317 following sense: on each choice list, reporting the urn reflecting one’s true upper or lower prob-
 318 ability is in one’s best interest—it weakly dominates any other report in the respective choice
 319 list. Hence asking a subject for an urn $[r, 1 - b]$ such that, in each of the branches on the 2D
 320 choice list, she prefers the option she would receive under the mechanism incentivises reporting
 321 a MPI. Since precise probabilities (and SEU) are a special case of multiple priors (respectively,
 322 α -maxmin EU), this mechanism functions equally for Bayesian decision makers, who are incen-
 323 tivated to report their precise probabilities.

324 Note that despite the higher complexity involved in eliciting probability intervals as opposed
 325 to precise probability values, this incentive mechanism is as parsimonious as standard choice
 326 lists for MPs. In the latter, MPs are determined by the switching point, i.e. the maximum
 327 probability for which the subject prefers the bet over the lottery with that probability. Similarly,
 328 the proposed probability-interval incentive mechanism only asks for a single point, which is the
 329 switching point on each branch of the 2D choice list. In standard MP choice lists, the switching
 330 point determines the preferences in the rest of the choice list by stochastic dominance. Similarly
 331 here, the elicited point determines the other preferences in the 2D choice list according to
 332 following property, which can be thought of as a probability-interval analogue of stochastic
 333 dominance.

334 **Definition 1** (Lower Stochastic Dominance). For all urns $[r, 1 - b]$ and $[r', 1 - b]$, $(z, [r, 1 - b], 0) \prec$
 335 $(z, [r', 1 - b], 0)$ whenever $r < r'$.

336 This says that, between ILs $(z, [r, 1 - b], 0)$ and $(z, [r', 1 - b], 0)$ corresponding to bets on red from
 337 urns with the same minimum proportion of blue balls, the decision maker prefers the prospect
 338 where the minimum proportion of red balls is higher. It is straightforward to check that this is
 339 the property behind preference patterns (6) and (7).

340

341 2.6. Implementation 2: Binary-choice procedure

342 Our second elicitation technique, which we implement in tandem with 2D choice lists (Section
 343 3.4), is a ‘bisection-style’ binary-choice procedure for identifying the MPI. Here we set out its
 344 general principles; full details are provided in Appendix A.2. The logic can again be illustrated
 345 on Figure 1. The space of interval-valued urns is divided into four preference-defined areas,
 346 summarised in Table 1. The procedure is based on the following observation.

347 **Proposition 2.** *Suppose preferences are represented according to (1) with $\alpha > \frac{1}{2}$, and let E be*
 348 *an event.*

Name	Preferences	Colour (in Figure 1)
R-B	$(z, [x, y], 0) \succeq (z, E, 0)$ $(0, [x, y], z) \succeq (0, E, z)$	Red & Blue
W	$(z, [x, y], 0) \preceq (z, E, 0)$ $(0, [x, y], z) \preceq (0, E, z)$	White (neither Red nor Blue)
R	$(z, [x, y], 0) \succeq (z, E, 0)$ $(0, [x, y], z) \preceq (0, E, z)$	Red
B	$(z, [x, y], 0) \preceq (z, E, 0)$ $(0, [x, y], z) \succeq (0, E, z)$	Blue

Table 1: Preference-based division of \mathcal{I}

349 a. For any point $[x, y]$ in the R-B region (i.e. such that the corresponding preferences in Table
350 1 hold, for E), $\underline{p}(E) \leq x$ and $\bar{p}(E) \geq y$. Moreover, for any point $[x, y]$ in the W region,
351 $\underline{p}(E) \geq x$ and $\bar{p}(E) \leq y$.

352 b. For any point $[x, y]$ in the R region (i.e. such that the corresponding preferences in Table
353 1 hold, for E), every $[x', y']$ with $x' \geq x$ and $y' \geq y$ is also in R. Moreover, for any point
354 $[x, y]$ in the B region, every $[x', y']$ with $x' \leq x$ and $y' \leq y$ is also in B.

355 It follows from part a. that if the experimenter has found a point (interval-valued urn)
356 $[x_{RB}, y_{RB}]$ in the R-B region (i.e. such that there is the preference pattern in Table 1, row 1),
357 and a point $[x_W, y_W]$ in the W region, then the MPI is contained in the ‘box generated’ by these
358 points, i.e. it is in the set $\{[x, y] : x_W \leq x \leq x_{RB}, y_{RB} \leq y \leq y_W\}$. The procedure works by
359 searching the smallest such generated box for further points in R-B or W, in order to ‘reduce’
360 the size of the boxes and hence ‘home into’ the MPI. In this sense, it is analogous to the bisection
361 procedure for MPs, where preferences indicate that the MP is in a particular interval, and the
362 procedure searches to reduce the width of that interval.

363 Note that a similar result to Proposition 2 a. for the R and B regions does not hold.⁵
364 However, by part b. it can be concluded, for any point (interval-valued urn) $[x, y]$ in R, that
365 every point North-East of $[x, y]$ is also in R, and similarly for a point in B. So, if the experimenter
366 has just discovered a point in R (i.e. the elicited preferences for the relevant urn are as set out
367 in Table 1), then, to seek a point in R-B or W, she needs to look South-West of this point; and
368 analogously for points in B. The procedure works, when at a point in R and B, by performing
369 a bisection along one-dimensional cuts of the space \mathcal{I} guided by this observation, until a point
370 in R-B or W is found. Then the procedure between the closest R-B and W points re-applies, in
371 an attempt to generate a ‘smaller’ box. Details are available in Appendix A.2.

372 One final important property of the procedure used is an in-built ‘precision bias’. Whenever

⁵This can be seen by considering the horizontal and vertical bold lines in Figure 1 to define four quadrants, and by noting that there are both red and blue areas in the upper left-hand and lower right-hand quadrants.

	Type of Source of uncertainty	Treatments	# Events	Elicited	# subjects
EXP 1	Minimum winter temperature ($^{\circ}\text{C}$)	Where: Paris; Sydney	4	MPIs, MPs (Paris)	80
EXP 2	Entrance exam grade (/20)	Exam: Maths, Contraction	5	MPIs	52

Table 2: Summary of studies

no point in the areas R-B or W has been found, and hence at a stage when it is searching for points in these areas, the procedure deliberately moves closer to the space of precise urns (the 45° line in Figure 1); again, see Appendix A.2 for details. In this way, if there is any misclassification of subjects, the tendency would be for the procedure to represent them as more precise than they actually are.

3. Experimental Methods

We carried out two experiments in which we used our method to elicit upper and lower CDFs for various sources of uncertainty. Both experiments involved two comparable yet different sources of uncertainty (Table 2). EXP 1 implemented a faster elicitation, eliciting probability intervals for fewer points per source. This left time for standard matching probability elicitation for the same events, which yields insights into the α -maxmin EU model (see Section 4.5). By contrast, EXP 2 implemented a slower upper and lower CDF elicitation, eliciting more points per source. It also involved an omnibus confirmation screen, allowing subjects to confirm or revise all choices concerning events in a source, after elicitation. In EXP 2, no MPs were elicited.

3.1. Subjects

132 subjects (undergraduate students) were recruited from two French academic institutions: 80 from university of Paris 1 for EXP 1 and 52 from HEC Paris Business School for EXP 2 (Table 2). Subjects' choices were collected through computer-based individual interviews that lasted about one hour in each of the two studies. Each individual interview started with a video presentation of the experimental instructions, followed by comprehension questions and one training MPI elicitation task (on an event not involved in the ensuing experiment).⁶ Appendix C reports the typical screenshots faced by subjects for the tasks. In both experiments, subjects were told that there were no right or wrong answers, and that they could ask any question regarding the experiment. Differences in experimental instructions between the experiments are explained in the sequel.

⁶The video presentations are available upon request.

398 **3.2. Sources of uncertainty**

399 Each experiment involved two comparable sources of uncertainty, with one treatment for each
400 source (Table 2). The type of source in EXP 1 was the minimum daily temperature over the
401 previous November–March period; the sources differed in the city whose temperature was of
402 interest—Paris, where the experiment was carried out, and Sydney. EXP 2 involved marks
403 in two of the previous year’s entrance exams for admission at undergraduate level through
404 the ‘ECS’ and ‘ECE’ entrance streams to a prominent French business school, HEC Paris.
405 The subjects in the experiment—students admitted at this level at the school—had sat the
406 exams either the same year or the previous year. The sources differed in the exam considered:
407 the probability and statistics exam (officially called ‘Mathématiques II’), which is generally
408 considered to be ‘objectively marked’, and the ‘Contraction’ exam—a summary of a philosophical
409 or literary text—whose marking is considered more ‘subjective’, ‘random’ and ‘unpredictable’
410 among candidates and students. Indeed, the marks in the latter exam have higher variance.⁷

411 Each source of uncertainty involves a variable (temperature in °C, mark out of 20), the
412 aim was to elicit subjects’ multiple prior beliefs—in the form of the generated upper and lower
413 CDFs (Section 2.2)—over the variable. In each experiment, the subject chose a number at the
414 beginning of the experiment⁸ which identified, according to a spreadsheet to which the subject
415 would only have access at the end of the experiment, a random day D between 1 November and
416 31 March of the previous Winter (in EXP 1), and a random candidate C for entry to HEC Paris
417 in the previous Spring (in EXP 2). We estimated upper and lower CDFs by eliciting subjects’
418 upper and lower probabilities for cumulative events, i.e. events E_{t_i} of the form: “the minimum
419 temperature on day D in Paris (resp. Sydney) was less than or equal to t_i ”, or “candidate C
420 obtained a mark less than or equal to t_i in the Maths (resp. Contraction) exam” (Section 2.2),
421 for various fixed values of t_i given in Table 3.⁹

422 Note that the events used pertained to time periods several months before subjects par-
423 ticipated in the experiment (temperature the previous Winter, for subjects taking part in the
424 experiment in Spring; exams sat the previous Spring, for subjects taking part in the experiment
425 in Autumn). Moreover, there is a natural difference in the familiarity with or predictability of
426 the sources involved in each experiment—with Paris’s weather being more familiar to Paris sub-
427 jects than Sydney’s, and Maths considered a more predictable exam than Contraction. Finally,
428 we had access to the real data for all the sources, which were used for incentivisation (Section
429 3.5).¹⁰

⁷The variance of marks for Maths is 3.77, where it is 9.92 for Contraction.

⁸They chose a number between 1 and 150 in EXP1 (the number of days in the period considered), 1 and 456 (the number of candidates) in EXP 2.

⁹For each source in EXP 1, we chose temperature values close to the 10%, 33%, 66% and 90% percentiles of the true distribution. For EXP 2, we used the same values for both sources (Maths and Contraction), picked so they would seem to reasonably scan the range and correspond to comparable points in the true distribution over Contraction scores, where they were at the 3% 15% 33% 68% and 86% percentiles. They were at the 0%, 0%, 2%, 21% and 60% of the true distribution of Maths scores.

¹⁰For the weather, the data source was Météo France (Paris Orly meteofrance.fr) and the Australian Bureau

	Treatment	Events $E_{t_i} = \{t' \in T : t' \leq t_i\}$ for t_i :
EXP 1	Paris	-2, 2, 5, 8
	Sydney	15, 18, 20, 22
EXP 2	Maths	7, 10, 12, 15, 17
	Contraction	7, 10, 12, 15, 17

Table 3: Sources of uncertainty and events

430 **3.3. Choice Tasks**

431 **EXP 1**

432 EXP 1 consisted of three blocks of tasks. Each of the first two blocks concerned a single source
433 (Paris or Sydney), and involved the elicitation of the upper and lower probabilities for each of
434 the events in the source (Table 3). The order of these two blocks was randomized. In each block,
435 the subject first declared, in an non-incentivized manner and using a scrollbar, her estimated
436 maximum and minimum values for the minimum temperature on the unidentified day selected.
437 This is standard procedure in expert elicitation for unbounded sources, aimed at combatting
438 anchoring bias (Morgan, 2014), and played no role in our elicitation. Then the elicitation
439 procedure set out in Section 2 and implemented as described in Section 3.4 was applied for
440 each event in the source. Within each block, the two extreme events (i.e. lowest and highest
441 temperature points) were asked first, in a random order, followed by the other two events, in a
442 random order.

443 The final block involved the elicitation of MPs for the events in Paris treatment. MPs
444 were elicited for each event E_{t_i} in this source and its complement $E_{t_i}^c$ (Table 3). The order of
445 elicitations was randomized in this block.

446 **EXP 2**

447 EXP 2 consisted of two blocks of tasks, corresponding to the first two blocks of EXP 1. Each
448 of the blocks concerned a single source (Maths or Contraction), and involved the elicitation of
449 the upper and lower probabilities for each of the events in the source (Table 3). The order of
450 the blocks was randomized. In each block, the elicitation procedure set out in Section 2 and
451 implemented as described in Section 3.4 was applied for each event in the source. The order of
452 events in each block was randomized. Each block ended with an omnibus confirmation screen,
453 in which the interval-valued urns elicited for each of the events in the source were displayed and
454 graphed, and the subject was given the opportunity to go back and modify any of her responses
455 for the events in the source (Section 3.4). This screen, the sources and the larger number of

of Metereology (Sydney Observatory Hill bom.gov.au); for the marks, they were provided by HEC admission services.

456 events elicited per source were the central differences with respect to EXP 1.

457 3.4. Elicitation procedures

458 Upper and lower probabilities in EXP 1 and EXP 2

459 Our elicitation procedure follows the general hybrid structure adopted by Abdellaoui et al. (2021)
460 for MP elicitation, under which a bisection procedure is used to aid subjects to fill in responses
461 on a choice list, which they then confirm or modify. For each event E_{t_i} (Table 3), we first
462 applied the binary-choice procedure set out in Section 2.6 and Appendix A.2. Each step of the
463 procedure involved an event E_{t_i} and a 100-ball urn with a specified minimum number of blue
464 and red balls, where nothing was known about the colour of the remaining balls. At each step,
465 two choices were elicited from subjects: their choice in the decision between the bet on the event
466 E_{t_i} and the bet on the next ball drawn from the urn being red, and their choice in the decision
467 between the bet on $E_{t_i}^c$ (or against E_{t_i}) and the bet on the next ball drawn from the same
468 urn being blue. (Details on the display are provided in Appendix C.) The urn proposed in the
469 next step depended on the preferences elicited in the previous step according to the procedure
470 (Section 2.6 and Appendix A.2). The subjective probability interval for E_{t_i} elicited at the end of
471 the procedure is deduced from the preferences over such bets, as specified in the cited sections.
472 The procedure continued until the interval was estimated to a precision of 0.15 if it was not
473 degenerate, 0.05 if it was degenerate (i.e. corresponding to a precise probability), or up to 12
474 steps, whichever came first.

475 At the end of the binary-choice procedure, the ‘confirmation’ 2D-choice list described in
476 Section 2.5 was displayed for verification. Although the ‘space of choices’ to be confirmed is the
477 two-dimensional ‘cross’ in Figure 1, we implemented it via a one-dimensional scrollbar-based
478 display with two cursors (see Appendix C). The cursors specified the minimum number of red
479 and blue balls respectively, and hence together determined an interval-valued urn. They were
480 initially set at the values determined by the binary-choice procedure. To confirm the whole 2D-
481 choice list, the subject had to scan all the associated choices. When moving the red cursor, the
482 blue cursor remained fixed at the pre-specified value. This accentuates the separate nature of
483 the cursors, which cannot be moved in tandem. By moving the red cursor, the subject scanned
484 all the urns with the same minimum number of blue balls but differing minimum numbers of red
485 balls, i.e. the choices represented by the bold red horizontal line in Figure 1. During this scan
486 the corresponding choices between the bets on the event E_{t_i} and the bets on red from the urn
487 were displayed, with the ‘chosen bet’, specified as in Section 2.5, being indicated (i.e. the bet on
488 the urn when there are more red balls than the provisionally elicited point; the bet on the event
489 otherwise). The subject also had to scan the choices associated with moving the blue cursor—
490 there, the red cursor (and hence minimum number of red balls) was held fixed. When moving
491 the blue cursor, the choices between the bets on $E_{t_i}^c$ and the bets on blue from the urn were
492 displayed (with the corresponding choice, following the logic in Section 2.5). This corresponds
493 to scanning the choices represented by the bold blue vertical line in Figure 1. By clicking on the

494 appropriate bet (on the event or the urn) in any of the displayed choices, subjects could revise
495 their reported preferences, hence modifying the specified position of the fixed cursors (and the
496 associated provisionally elicited point). After such modifications, subjects had to reconfirm all
497 of the associated choices, by moving one and then the other cursor, before moving on to the next
498 stage of the experiment. The precision of the scrollbar, and hence subject responses, was to the
499 nearest 0.01 (to the precise minimum number of red and blue balls out of 100 respectively).

500 Omnibus confirmation screen in EXP 2

501 In EXP 2, after the procedure described above was completed for all the events in the source,
502 the subject was asked to confirm all the elicited values, and given the opportunity to modify
503 responses. The confirmation screen displayed the interval-valued urns elicited for the five events
504 in the source. Moreover, this information was summarized in a graph displaying the minimum
505 number of red and blue balls for each event (see Appendix C). Hovering the mouse over the
506 points on the graph caused the associated interval-valued urn to be highlighted. By clicking
507 on the point on the graph or the urn, the subject could access the corresponding two-cursor
508 scrollbar confirmation screen at the end of the binary-choice procedure for that event, where she
509 could change her choices in exactly the same way as set out above.

510 Matching probabilities in EXP 1

511 The MP of the bet on a given event was elicited through a two-step procedure, from which
512 our multiple prior elicitation procedure was inspired. First, a candidate MP was determined
513 through a bisection process (Abdellaoui et al., 2008) that consisted in a chained sequence of
514 binary choices between the bet on the event and an urn whose composition was fully known.
515 Starting with a binary choice between $(z, [\frac{1}{2}, \frac{1}{2}], 0)$ and $(z, E, 0)$, it then asks a binary choice with
516 the midpoint of the lower (respectively upper) interval $[0, \frac{1}{2}]$ (resp. $[\frac{1}{2}, 1]$) whenever the subject
517 chooses the former (resp. latter) option, and so on. The displays used were similar to those
518 described above. Then the complete confirmation (one-dimensional, single cursor) scrollbar-
519 based choice list, filled in according to the prior bisection choices, was displayed for verification.
520 The precision of the elicited MP was to the nearest 0.05.

521 3.5. Incentivizing subjects

522 Participants in all studies received a flat payment of €10. Additionally, a random incentive
523 system was implemented, which was entirely analogous to those standardly used to implement
524 elicitation of matching probabilities. As noted above, after the presentation of the instructions
525 and before the beginning of the experiment, the subject chose a number from a given range,
526 which identified an individual case of the variable of interest (the day, if the source was minimum
527 temperature; the candidate, if the source was the mark). The exact case identified was specified
528 according to a spreadsheet that would only be revealed at the end of the experiment. This is
529 in concordance with the approach set out by Johnson et al. (2021), who argue that it reduces

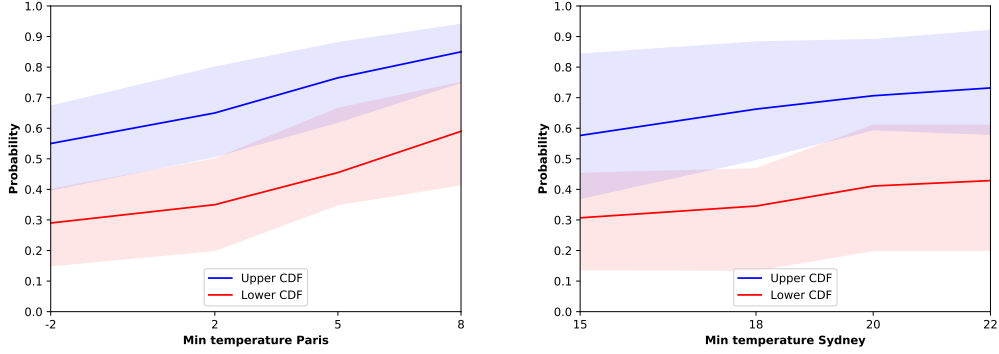
530 hedging motivations, given the well-known fact that ambiguity models are indifferent to ex ante
531 hedging. At the end of the experiment, a choice list (a 2D-choice list or MP-choice list in
532 EXP 1; a 2D-choice list in EXP 2) and choice on it were chosen at random by the computer.
533 The subject was then paid according to the decision she had made on that choice. If she had
534 chosen, say, the bet on the event that the minimum temperature in Paris is less than or equal
535 to 2°C, then the day which she chose was revealed, as well as daily temperature data for the
536 November–March period, and she won if the minimum temperature on that day was indeed 2°C
537 or less; if not, she lost. If she had chosen the urn, then she composed the appropriate urn—she
538 counted the specified minimum numbers of red and blue balls, with the remaining balls coming
539 from pre-constructed Ellsberg urns (of unknown composition). Then a ball was drawn from the
540 constructed urn, and she was paid according to whether she bet on the color of that ball or not.
541 All bets yielded 20€ if won, and nothing otherwise.

542 4. Results

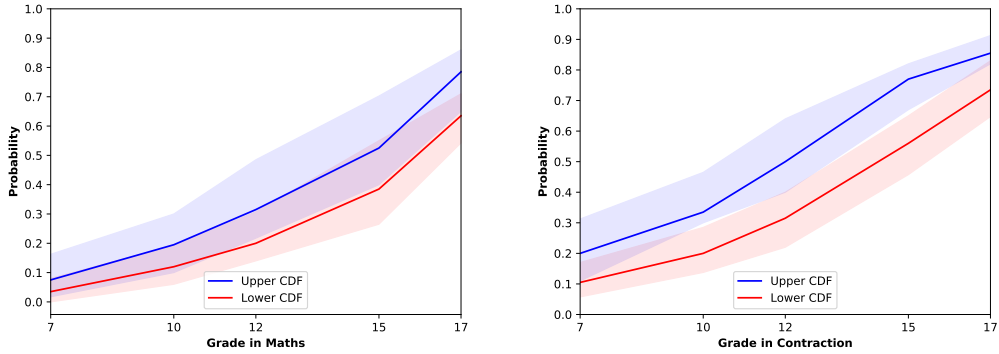
543 4.1. Descriptive Statistics and Performance

544 Figure 2 plots the median, 25% and 75% quantile upper and lower CDFs for all elicited events and
545 both experiments (see Tables 5–6 in Appendix B.1 for basic descriptive statistics). This Figure
546 already gives some early indications about our results, and the performance of our elicitation
547 method.

548 First of all, it reports ‘well-behaved’ upper and lower CDFs with probabilities differing across
549 subjects and events—thus suggesting the consistency of the method. The elicited points for both
550 upper and lower CDFs were also consistent across the successive steps of the elicitation procedure:
551 the binary-choice procedure and the confirmation 2D choice list (Tables 5–8, Appendix B.1).
552 Moreover, it suggests that, in the aggregate, upper and lower CDFs are increasing, as they should
553 be. Figure 3 plots descriptive statistics of the empirical distribution of individual Kendall τ_b
554 rank correlation coefficients between the size of events and the upper (resp. lower) probabilities
555 or MPs elicited for each source (see also Table 10, Appendix B.1). As is clear from the Figure,
556 the median Kendall τ_b is far greater than 0 for all sources, pointing to increasing upper and
557 lower CDFs. There is however a notable difference between EXP 1 and EXP 2. In EXP 2,
558 where subjects were given the opportunity to confirm all their replies on all the 2D choice lists
559 for a source (Section 3.4), CDFs were strictly increasing for the vast majority of subjects. In
560 EXP 1, where 2D choice lists were confirmed after consideration of the event and there was
561 no opportunity to reconfirm later, there were more violations of monotonicity. Comparison
562 of the (cognitively less demanding) MPs with upper and lower CDFs in the Paris treatment,
563 whose median Kendall ranks are similar (Figure 3a), suggests that such violations were not
564 unique to the elicitation method proposed here. As could have been expected, the frequency of
565 monotonicity violations appears to increase with the difficulty of the choice task, with the MP
566 task being arguably easier than that for probability-interval elicitation, and the task for Paris,



(a) EXP 1



(b) EXP 2

Figure 2: Median, 25% and 75% quantile ranges of upper and lower CDFs

567 the more familiar source for our subjects, being easier than that for Sydney.

568 4.2. Bayesian analysis

569 We also adopt a standard Bayesian approach, estimating hyperparameters for upper and lower
 570 CDFs using a MCMC procedure. We run estimations for each source under the assumption that
 571 upper and lower CDFs follow a (truncated) normal distribution, and under a Beta distribution
 572 (Table 14, Appendix B.2). As shown in Table 4, the Beta distribution has the best goodness of
 573 fit under both the AIC and BIC criteria for the sources in EXP 1, whereas the truncated Normal
 574 distribution performs better according to both criteria for the sources in EXP 2. Henceforth,
 575 we present the results under these distributions (the analyses under the other distributions are
 576 given in Appendix B.2).

577 Figure 4 plots 1000 MCMC samples for each of the upper and lower distributions, for each
 578 source. (Statistics on the distributions of parameters are given in Tables 15–22, Appendix
 579 B.2.) They suggest that the proposed elicitation technique supports parametric estimation of
 580 subjective probability intervals in the population, insofar as they chime with expectations given
 581 the nature of the events. For instance, they suggest that the dispersion of subjective upper and
 582 lower probabilities is larger for the temperature source (EXP 1) than the grade source (EXP 2),

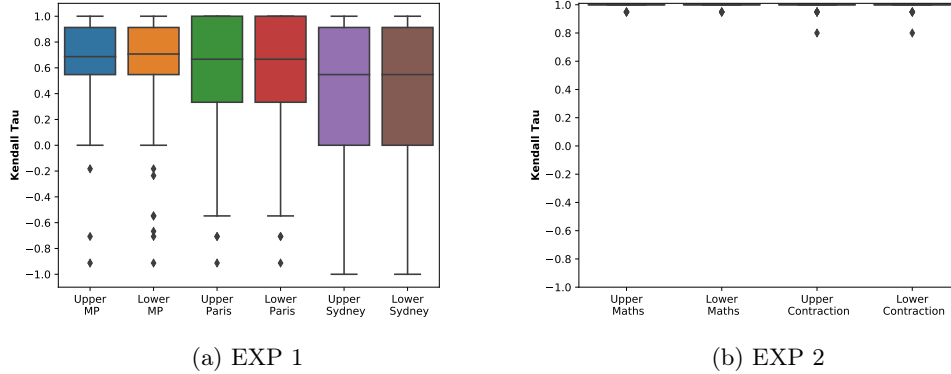


Figure 3: Individual-level Kendall τ_b for upper and lower CDFs for the sources for each experiment, and for the MPs in EXP 1, Paris treatment. In the case of MPs, the Lower MP is calculated using the MPs of the events E_t (and should be increasing in t), whereas the Kendall τ_b for upper MPs are calculated using one minus the MPs of complementary events E_t^c (which should be increasing with t).

Note: The Kendall τ_b is an indicator of ordinal association: the value 1 indicates that the CDFs or MPs are strictly increasing; 0 suggests that there is no association between the elicited probability and the size of the event; -1 indicates a strictly decreasing relationship between the two.

		Paris	Sydney	Mathematics	Contraction
AIC	normal	706.65	700.79	411.22	385.52
	beta	648.26	684.36	416.18	390.64
BIC	normal	711.42	705.56	415.12	389.42
	beta	653.02	689.12	420.08	394.54

Table 4: AIC and BIC under (truncated) normal and Beta specifications for CDFs (Table 14).

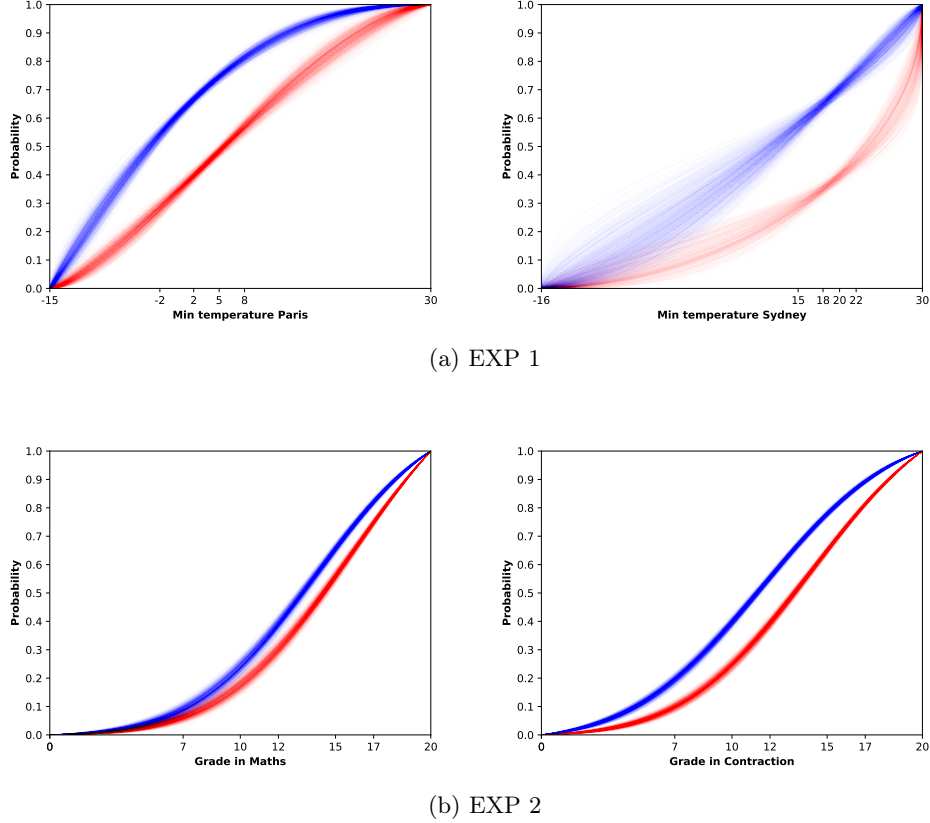


Figure 4: Bayesian estimation of lower and upper CDFs: plots of 1000 samples from MCMC. (Beta distribution for EXP 1; Truncated Normal distribution for EXP 2)

583 which could be related to the fact that all subjects in EXP 2 had sat both exams, and were very
 584 interested in the marking, several months before. Also, within EXP 1, there is more dispersion
 585 in the estimated distributions for Sydney than for Paris, as would be expected given the less
 586 familiar nature of the former source for Paris subjects.¹¹

587 4.3. Imprecision

588 Both the graphs of raw data (Figure 2) and those emerging from the Bayesian analysis (Figure
 589 4) suggest that subjects' beliefs are often *imprecise*: i.e. there is a gap between their upper
 590 and lower probabilities. Indeed, two-sided Kolmogorov-Smirnoff tests of the hypothesis that
 591 the median upper and lower CDFs are drawn from the same distribution reject the hypothesis
 592 for each source ($p < 0.0001$ in all cases), suggesting a gap between upper and lower CDFs.
 593 For further analysis, we define the following index. For an event E from a given source (e.g.
 594 minimum temperature in Paris), we say that a subject's *imprecision concerning E* is $\bar{p}(E) - \underline{p}(E)$,
 595 i.e. the width of her (elicited) probability interval for E . A subject's *Imprecision Index* for a
 596 source is defined to be her average imprecision across all elicited events in the source:

¹¹More precisely, it is clear from Tables 16 and 18 that the standard deviations of the parameters for the Paris source are lower than for Sydney.

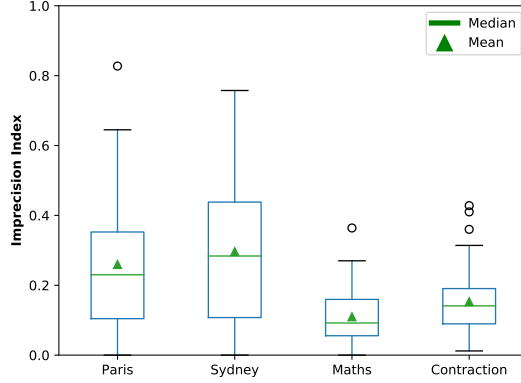


Figure 5: Imprecision Index (Eq. (8)) across sources in EXP 1 and EXP 2

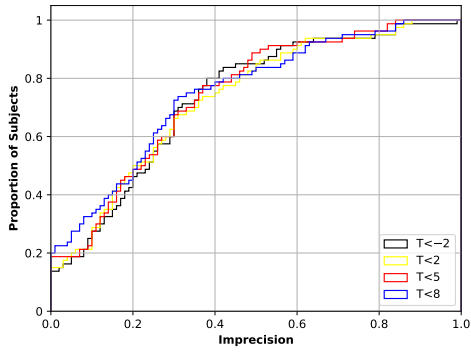
$$II = \frac{1}{n} \sum_{i=1}^n (\bar{p}(E_n) - \underline{p}(E_n)). \quad (8)$$

597 This clearly gives an indication of how imprecise the subject’s beliefs are, on average, for events
 598 in the source. Naturally, an SEU decision maker will assign precise probabilities to all events,
 599 and hence have an imprecision index of 0 (for all sources).

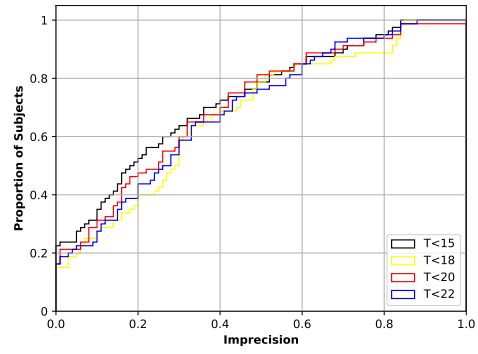
600 Figure 5 displays the mean, median, 25% and 75% quantile, and max and min Imprecision
 601 Indices across all sources in both experiments (see also Table 11, Appendix B.1). It clearly
 602 suggests a tendency towards imprecision, with mean and median Imprecision Indices greater
 603 than 0.1 for all sources. Two-sided binomial tests reject the hypothesis of equal probability for
 604 the Imprecision Index to be equal to vs. greater than 0 for each source ($p < 0.0001$ in all cases),
 605 with a clear majority of subjects—74 out of 80 in EXP 1, and 49 out of 52 in EXP 2—having
 606 strictly positive Imprecision Indices.

607 The general message of widespread imprecision is confirmed by data on the number of precise
 608 events—events for which the subject’s elicited upper and lower probabilities coincide (Table 12,
 609 Appendix B.1). Not more than around 5% of subjects gave precise probabilities for all events in
 610 a single source. Only 2 subjects (out of the 132 participating in both experiments) gave precise
 611 probabilities for all events elicited. The data in Table 12 also allows a check on the extent to
 612 which this imprecision could be driven by the binary-choice procedure, insofar as it gives the
 613 number of precise events after the binary-choice procedure and before the confirmation 2D choice
 614 list, as well as after confirmation. The general finding of few fully precise subjects holds both
 615 before and after the confirmation stage. Moreover, relatively few subjects change to fully precise
 616 probabilities for all events of the source (at most 3 out of 80, for Sydney in EXP 1), with several
 617 fully precise subjects introducing imprecision during the confirmation stage, especially in EXP
 618 2.

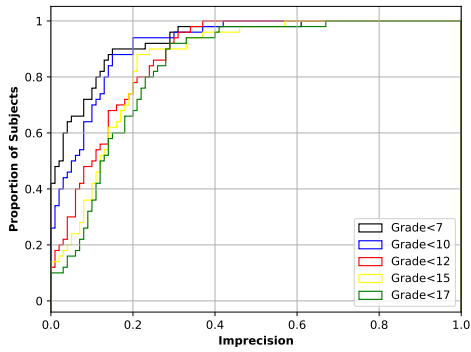
619 Delving further, we also investigate imprecision at the event level within sources. Figure 6
 620 plots CDFs of the imprecision for each elicited event in each of the experiments and sources,



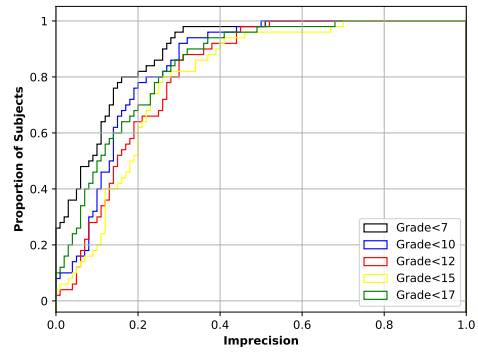
(a) Paris source, EXP 1



(b) Sydney source, EXP 1



(c) Maths source, EXP 2



(d) Contraction source, EXP 2

Figure 6: CDFs of Imprecision across subjects, for each elicited event

621 across subjects. One-way ANOVAs of the imprecision (dependent variable) against the event
 622 (factor) reject the null hypothesis of identical imprecision across all events for the sources in EXP
 623 2 ($p < 0.001$ for Maths; $p = 0.003$ for Contraction), whilst failing to reject it for the sources in
 624 EXP 1 (Table 13, Appendix B.1.1). This suggests not only that imprecision is widespread, but
 625 that imprecision may be event dependent within sources, as one would expect if some events
 626 are intuitively more uncertain than others. For instance, the least imprecise event in EXP 2
 627 involves, for both sources, the lowest grade, where many subjects are presumably more sure of
 628 their judgements.

629 In summary, the development of a method for eliciting multiple priors does not emerge from
 630 this analysis as devoid of relevance: rather, it reveals that, when given the possibility to ‘express’
 631 the imprecision implied by non-degenerate probability intervals, many subjects do, at least for
 632 the events considered here. Moreover, at least within some sources, the extent of imprecision
 633 may depend on the event.

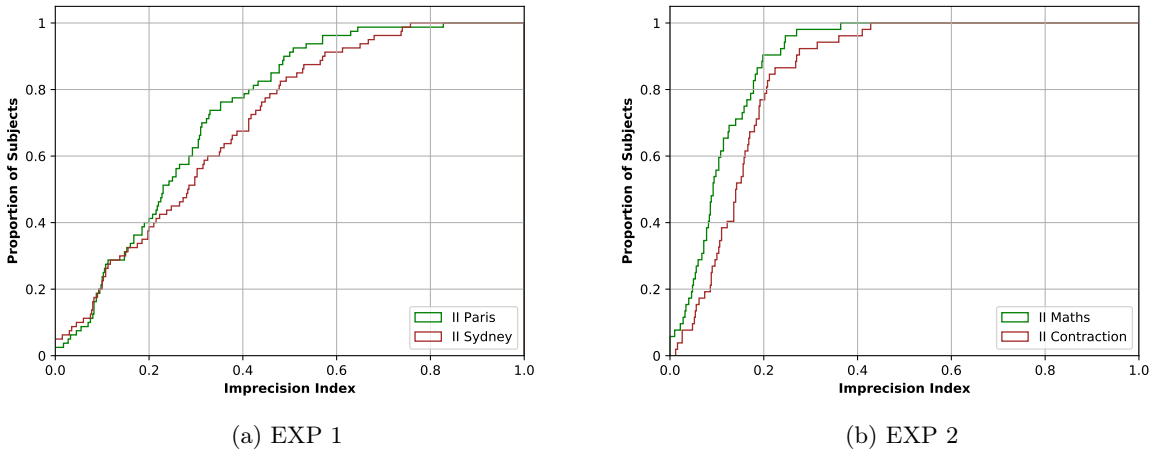


Figure 7: CDFs of Imprecision Index (Eq. (8))

634 4.4. Imprecision and familiarity

635 One reasonable hypothesis is that *ceteris paribus* subjects' beliefs are more imprecise concerning
 636 events with which they are less familiar, or about which they feel as if they have less knowledge.
 637 In terms of multiple priors models, this corresponds to the probability intervals for the events
 638 being wider. Since, as explained in Section 3, each of our experiments features two sources
 639 with which our subjects will typically have different levels of familiarity, or which they naturally
 640 consider as having different levels of predictability, a natural conjecture would be that imprecision
 641 would be larger for Sydney than Paris, and for the Contraction grade than the Maths one. After
 642 all, Paris subjects are less familiar with the weather in Sydney than that in Paris; and the
 643 Contraction exam is generally considered to be 'less predictable' than the Maths one (Section
 644 3.2).

645 Figure 7 plots the CDFs of the Imprecision Index defined above (Eq. (8)) across subjects, for
 646 the pair of sources in each experiment. A two-sided paired *t*-test barely fails to reject the null
 647 hypothesis of identical Imprecision Indices across the sources in EXP 1 ($p = 0.0895$), whilst it
 648 rejects it for EXP 2 ($p = 0.0016$). A two-sided Binomial test with null hypothesis that an equal
 649 number of subjects have larger Imprecision Index under one source than the other fails to reject
 650 the null hypothesis for EXP 1 ($p = 0.576$), but rejects it for EXP 2 ($p = 0.017$). Rerunning the
 651 latter test using the results of the Bayesian analysis—i.e. for each experiment and source, using
 652 the posterior distribution over parameters obtained from the Bayesian estimation to sample 1000
 653 tuples of parameters determining the upper and lower distributions, and computing, for each
 654 tuple, the Imprecision Index as defined in Eq. (8)—yields a rejection of the null hypothesis
 655 across the sources for both EXP 1 and EXP 2 ($p < 0.001$ in both cases). These findings confirm
 656 the expected relationship between imprecision and predictability in EXP 2: indeed, in Figure
 657 7, the CDF for Contraction—known as the less predictable exam—is entirely to the right of
 658 that for Math, indicating a larger Imprecision Index. They also point to a similar relationship

659 between imprecision and familiarity in EXP 1: again, with Figure 7 suggesting that CDFs for
 660 Paris are generally more precise.

661 That the expected relationship between imprecision and familiarity or predictability emerges
 662 can also be seen as providing further indirect evidence as to the solidity of the proposed elicitation
 663 method.

664 4.5. Matching probabilities and the α -maxmin EU mixture coefficient

665 Recall that EXP 1 contained a supplementary treatment in which the MPs were elicited for the
 666 Paris events E_{t_i} and $E_{t_i}^c$ for which probability intervals had been elicited (Table 2). Henceforth,
 667 we denote the MP of an event E by $MP(E)$. Under SEU, $MP(E_{t_i}) = 1 - MP(E_{t_i}^c) = p(E_{t_i})$,
 668 the subjective probability of E_{t_i} , for all E_{t_i} . So, as is well-known, comparing $MP(E_{t_i})$ and
 669 $1 - MP(E_{t_i}^c)$ provides an indication into the violation of SEU. Under the α -maxmin EU model
 670 (1), we have the following equations:

$$MP(E_{t_i}) = \alpha \underline{p}(E_{t_i}) + (1 - \alpha) \bar{p}(E_{t_i}) \quad (9)$$

$$1 - MP(E_{t_i}^c) = \alpha \bar{p}(E_{t_i}) + (1 - \alpha) \underline{p}(E_{t_i}) \quad (10)$$

671 Drawing on the elicited MPs and our elicitation of upper and lower probabilities, the equations
 672 (9) and (10) can be used to elicit the mixture coefficient α in the Hurwicz α -maxmin EU model.
 673 Under analysis using the raw data, the median α across subjects is 0.80 (Table 24, Appendix
 674 B.3). We also perform a Bayesian estimation of the α in tandem with the lower and upper CDFs,
 675 combining equations (9) and (10) and the MP data with our upper and lower CDF elicitation
 676 (see Appendix B.2, Tables 15 and 16). Figure 8 plots the distribution over α resulting from this
 677 estimation. The Bayesian mean for α is at 0.81, which is broadly consistent with the finding
 678 from the raw data. As discussed at more length in Section 5, this is, to our knowledge, the first
 679 direct choice-based elicitation of the α in the α -maxmin EU model that fully controls for the
 680 set of priors by eliciting the relevant information about them without making any assumption
 681 about their shape.

682

683 5. Discussion

684 The proposed method elicits non-degenerate and reasonable upper and lower CDFs. Our elicitations
 685 also show that imprecision—a gap between upper and lower probabilities—is widespread,
 686 with few subjects having precise probabilities for all events. Moreover, they bring out some
 687 determinants of imprecision. For some sources, the width of probability intervals may vary
 688 according to the event elicited; moreover, average imprecision decreases with the familiarity of
 689 the source of uncertainty, as one might expect. Finally, we draw on our probability interval

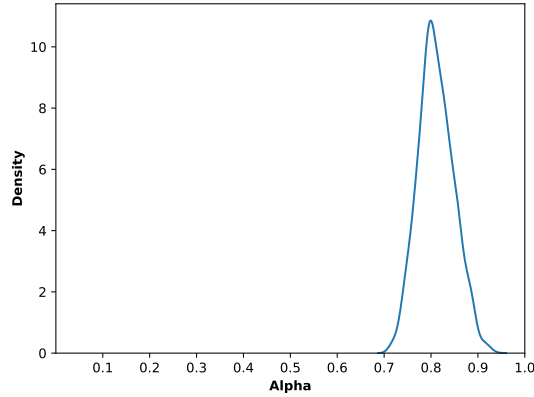


Figure 8: PDF of α from the Bayesian estimation (EXP 1, Paris treatment)

690 elicitation to elicit the mixture coefficient in the Hurwicz α -maxmin EU model—the first such
 691 elicitation, to our knowledge, to fully control for beliefs.

692 We now discuss the robustness of our procedure, some related literature, and some directions
 693 for future development.

694 **Robustness** Our approach has been presented in terms of the popular Hurwicz α -maxmin EU
 695 decision model (Section 2), which is doubtless one of the most general decision models in which
 696 the ‘belief component’ of the representation is just a set of priors. However, many of the central
 697 elements of the approach generalise to other models, including extensions building on sets of
 698 priors but weakening the linearity of the Hurwicz function form (1) to account for probability
 699 weighting, for instance (see Appendix A.3 for details). First of all, the notion of MPI remains
 700 well-defined for all such extensions, and the decision maker’s subjective probability interval is
 701 always a MPI. Though MPIs are not guaranteed to be unique for every conceivable extension
 702 of this sort, they are essentially unique for a family of reasonable extensions (Appendix A.3).
 703 Second, the 2D choice list incentivization mechanism only relies on the weak Lower Stochastic
 704 Dominance property of preferences (Definition 1, Section 2.5). Apart from the maxmax-EU
 705 model (i.e. (1) with $\alpha = 0$), which is very rarely found in subjects, this property is satisfied by
 706 any reasonable decision model generalising α -maxmin EU to allow for nonlinear dependence of
 707 preferences on upper and lower probabilities (Appendix A.3). In this sense, the 2D choice list
 708 incentivization mechanism is widely valid. Finally, whilst the binary-choice procedure relies on
 709 the strongest assumption made in Section 2—that $\alpha > \frac{1}{2}$ —there is independent evidence that
 710 this holds for most of our subjects (Appendix A.3). As noted in Section 3.4, it is the 2D choice
 711 list confirmation task that counts for incentivizing subjects’ choices, the binary-choice procedure
 712 playing the role of an aid to completing it.

713 The 2D choice list mechanism is incentive compatible whenever subjects treat the branches in
 714 the choice list in isolation from each other (Section 2.5); if this isolation assumption is violated,
 715 strategic choice may occur. Our implementation was designed to favour such isolation, notably

716 via the realisation of 2D choice lists by a single scrollbar with two cursors (Section 3.4 and
717 Appendix C). Visually very different from Figure 1, this presentation promotes considering
718 each branch in isolation and is less suggestive of strategic opportunities of changing the choice
719 lists by ‘moving around the 2D space’. Notwithstanding this, the extent to which isolation
720 holds in our experiments is ultimately an empirical question, and we treat it as such. On this
721 front, the proposed elicitation method has the advantage that strategic reasoning leads to easily
722 recognizable choice patterns. As discussed in Appendix A.3.2, for a subject represented by (1)
723 with $\alpha \in (0, 1)$ (and any set of priors), her optimal response to the 2D choice list task when
724 reasoning strategically is one of the points $[0, 0], [0, 1], [1, 1]$: i.e., one of the vertices of the space
725 of interval-valued urns in Figure 1. Examining the concentration of responses at the vertices thus
726 provides insight into the extent of strategic reasoning in our subject pool. Our data suggests that
727 it is very limited: in EXP 1, only one subject (out of 80) reported vertex points for more than
728 half of the elicited events,¹² whereas for no subject in EXP 2 were more than half of the elicited
729 points among the vertices (Table 9, Appendix B.1; see Appendix A.3.2 for further details).

730 **Related literature** Our elicitation method relates to existing experimental and theoretical
731 literature on multiple prior models, and the α -maxmin EU model in particular. Part of this
732 literature is concerned with testing such models, or comparing them to others (e.g. Hey et al.,
733 2010; Baillon and Bleichrodt, 2015); by contrast, the aim here is to elicit probability intervals
734 in the context of a fairly general multiple prior model. Similarly, there is a literature studying
735 willingness to bet on objectively-given probability intervals based on interval-valued urns (e.g.
736 Baillon et al. (2012); Chew et al. (2017)) using matching probabilities or certainty equivalents.
737 The present paper, by contrast, uses such urns in the context of the different task of eliciting
738 subjective probability intervals.

739 On the theory side, the challenge of incentive-compatible elicitation of multiple prior beliefs
740 under α -maxmin EU is related to identification issues with this model, arising from the fact that
741 different pairs of mixture coefficient α and sets of priors can represent the same preferences over
742 (Savage or Anscombe-Aumann) acts. Proposed approaches to this challenge include pinning
743 down the set of priors using ‘unambiguous preferences’ (Ghirardato et al., 2004), though this
744 has problems in finite state spaces (Eichberger et al., 2011), or enriching the state space to
745 include an infinite product structure and invoking symmetry axioms (Klibanoff et al., 2021). An-
746 other line of attack concentrates on special cases of the α -maxmin EU model, notably involving
747 some form of probabilistic sophistication, i.e. the assumption that there are precise probabilis-
748 tic beliefs which completely determine the contributions of events to preferences (Machina and
749 Schmeidler, 1992; Chew and Sagi, 2006). Working with a rich state space à la Savage, Gul and
750 Pesendorfer (2014, 2015) obtain a unique identification of α and the set of priors whenever the
751 latter is generated as the set of extensions of a precise probability measure on a subalgebra of
752 events. (Grant et al. (2019) have extended this approach beyond the assumption of linearity

¹²He / she reported 3 points out of 4 as $[0, 1]$, for both sources.

753 in upper and lower probabilities built into α -maxmin EU; see Appendix A.) Chateauneuf et al.
754 (2007) obtain a unique identification of α and the set of priors whenever the latter is gener-
755 ated from a precise probability measure via ε -contamination, i.e. mixture with the set of all
756 probability measures. As stated in the Introduction, we specifically avoid the sort of proba-
757 bilistic sophistication assumption behind these approaches, motivated by the observation that
758 such assumptions are inadmissible precisely in those situations where multiple prior beliefs are
759 most relevant. Indeed, our procedure takes a different approach, based on interval lotteries,
760 with no need for specific richness assumptions on the state space, probabilistic sophistication,
761 or any other (non-standard) assumptions on the set of priors. Hill (2021) provides an axiomatic
762 foundation for α -maxmin EU and extensions by generalising the elicitation principle used in the
763 present paper.

764 On the experimental front, there is a small literature dealing with incentive-compatible
765 elicitation of multiple priors. One family of approaches purport to elicit multiple priors as the
766 support of second-order beliefs, represented as a measure over the space of probability measures.
767 Beyond the assumption of second-order beliefs, which is foreign to the original multiple prior
768 models (Gilboa and Schmeidler, 1989; Bewley, 2002; Ghirardato et al., 2004), these often make
769 further assumptions about the role of these second-order beliefs in choice. For instance, Qiu and
770 Weitzel (2016) elicit subjects' distributions over the matching probabilities of other participants
771 in the experiment, and purport to deduce subjects' own second-order beliefs from these, relying
772 on the assumption that a subject's opinions about others' matching probabilities coincides with
773 the uncertainty surrounding her own assessment. In a theoretical paper, Karni (2020) develops
774 an ingenious incentive-compatible mechanism for eliciting second-order beliefs and the associated
775 set of priors (as the support), relying on a three-period setup. The mechanism assumes that the
776 subject's second-order beliefs coincide with her beliefs about what she will believe in the interim
777 period. As made clear above, our method relies on no assumptions beyond the α -maxmin EU
778 model (or appropriate weakenings thereof), and in particular there is no role for second-order
779 beliefs or assumptions on how they relate to other beliefs.

780 Another family of approaches draws on the theoretical literature discussed above, and in
781 particular on the probabilistically sophisticated special case studied by Chateauneuf et al. (2007),
782 where the subject's set of priors is generated as the ε -contamination of a single probability
783 measure with the space of all priors.¹³ Dimmock et al. (2015); Baillon et al. (2018b,a) use
784 elicitation of standard MPs (in the case of the last paper, certainty equivalents) to estimate
785 'ambiguity indices', from which one can back out the mixture coefficient α and the parameters
786 of the ε -contaminated set of priors. As noted previously, multiple prior decision models come
787 to the fore in situations where preferences cannot be reasonably assumed to be generated from
788 precise probabilities, and our elicitation technique was specifically designed to be independent
789 of the assumption of probabilistic sophistication for this reason. Moreover, our data provides

¹³Formally, the assumption is that the set of priors $\mathcal{C} = \{(1 - \varepsilon)p + \varepsilon\Delta\}$, where Δ is the space of all probability measures, p is an element of Δ and $\varepsilon \in [0, 1]$.

790 empirical insight into the aforementioned probabilistic sophistication assumption. In particular,
791 this assumption implies that the imprecision (in the sense of Section 4.3) is the same for all
792 events.¹⁴ As noted in Section 4.3 (see also Table 13, Appendix B.1.1), our observations reject
793 this equality for the sources in EXP 2, though not for the sources in EXP 1. This suggests that
794 there are sources for which their method’s underlying assumption does not hold. That said, it
795 may be viable on some sources; indeed, our data indicate that the Paris source in EXP 1 may be
796 one such source. And in fact, we can estimate the ambiguity indices used in the aforementioned
797 papers on the basis of the data from our study (EXP 1, Paris treatment) under their assumption
798 about the set of priors,¹⁵ and find, for instance, that they yield the value 0.82 for the mixture
799 coefficient α —which, reassuringly, is close to the Bayesian and raw estimates reported in Section
800 4.5. So not only is our elicitation method more robust, insofar as it applies in situations where
801 the assumptions underlying their approach do not hold, it can evaluate precisely in which cases
802 they do hold; in those cases, their approach, implemented on our data, gives the same result as
803 our ‘direct’ elicitation.

804 Going beyond the lab, there is a large and growing literature on elicitation of multiple priors
805 or imprecise probabilities in a range of disciplines, from economics to climate science. All such
806 elicitation exercises of which we are aware use stated probability intervals, and as such are not in-
807 centive compatible. For instance, Giustinelli et al. (2021) elicit beliefs on dementia and long-term
808 care decisions in a large-scale representative survey (over 1000 subjects), allowing stated prob-
809 abilities to be interval-valued. Consistently with our results (Section 4.3), they find widespread
810 imprecision. They argue forcefully for the importance of probability-interval elicitation for re-
811 ducing survey bias and understanding attitudes to and behavior in the face of high-uncertainty
812 events, such as whether one will develop dementia and whether one should insure against it. In
813 another approach, in another domain, Kriegler et al. (2009) elicit beliefs of selected scientists
814 (around 50 subjects) concerning climate tipping points, allowing participants to state probability
815 intervals for these (notoriously uncertain) events. Such expert elicitations, which involve often
816 time-consuming and individualised sessions with selected experts, have emerged as a central
817 tool for managing complex uncertainties (Morgan, 2014). Though they have traditionally aimed
818 at eliciting precise probabilities, Kriegler et al. (2009) shows that imprecision is widespread for
819 some events, and hence once again argue for the relevance of probability-interval elicitation.

¹⁴If the set of priors is as defined in footnote 13, then, for any E , $(1 - \varepsilon)p(E) \in [0, 1 - \varepsilon]$, so the probability interval for event E is $[(1 - \varepsilon)p(E), (1 - \varepsilon)p(E) + \varepsilon]$, and hence the event has imprecision ε .

¹⁵Specifically, Baillon et al. (2018b) propose the average of $1 - MP(E) - MP(E^c)$ over a selection of events as their measure of the ‘ambiguity aversion index’ b . The average for the events elicited here can be deduced directly from Table 25 (Appendix B.4), as around 0.16. On the other hand, under (1) with the specified form for the set of priors (see footnote 13), their ‘a-insensitivity index’ $a = \varepsilon$. Under such sets of priors, as noted in footnote 14, every E has imprecision ε . The average imprecision, as measured by the Imprecision Index (Table 11), thus gives an estimate of their a : it is around 0.25. The mixture coefficient α is related to these indices by $\alpha = \frac{1}{2} (\frac{b}{a} + 1)$ (Baillon et al., 2018b,a), yielding the value in the text.

820 **Future Directions** Two leitmotifs emerge from the literature review. On the one hand, our
821 results are consistent with existing studies suggesting that imprecision is widespread for some
822 events. However, based as they are on an incentive-compatible, choice-based and theoretically
823 robust elicitation method, our results are less open to criticisms of existing studies pointing to
824 a lack of incentive compatibility or the reliance on a specific model. On the other hand, as
825 we saw on the ε -contamination example, our method can be used to evaluate the assumptions
826 behind—and hence the effectiveness of—existing methods.

827 The latter point suggests one direction for future research. As noted, stated probability
828 intervals are typically used in large-scale surveys (such as Giustinelli et al. 2021), but how close, or
829 far, are subjects’ stated probability intervals from their actual multiple prior beliefs? Our method
830 can be used to provide insight into this question, for instance by eliciting probability intervals
831 with both the proposed method and stated procedures for the same subjects and on the same
832 set of events, and comparing the results. Although the ‘test’ events should involve uncertainty
833 that resolves in a reasonable timescale for payment of the incentives—so the method can be
834 applied in an incentivised fashion—the conclusions of such a comparison may be extrapolated to
835 situations where incentive payment is infeasible, for instance if the events of interest are too far
836 in the future (e.g. number of global pandemics before 2100) or counterfactual (e.g. how many
837 pandemics would there be if 60% of original wild habitats had been protected). As such, our
838 method can be used to corroborate, refine, correct and choose between existing stated approaches.

839 Moreover, although our method was developed with the aim of demonstrating the possibility
840 of choice-based incentive-compatible probability-interval elicitation, future research could op-
841 erationalise simpler, parametrised versions, with fewer choice questions, which would be more
842 implementable in field studies. Large-scale surveys often use choice tasks without necessar-
843 ily incentivising them (e.g. Falk et al., 2018), and questions formulated in terms of bets may
844 trigger different cognitive mechanisms to those formulated in terms of judged probabilities.
845 Our method could thus lay the foundations of a bet-based approach to add to the arsenal of
846 probability-interval elicitation procedures used in practice.

847 Finally, analogous possibilities exist for expert elicitation exercises, of the sort cited above.
848 Compared to survey studies, these typically involve fewer subjects, with each spending more
849 time; the flip side is that more precision is desired of the elicitation at the individual level. EXP
850 2 suggests that our method may provide the appropriate individual-level probability-interval
851 elicitation, whilst having theoretically well-founded incentive-compatibility properties. Prob-
852 ability elicitation exercises in decision analysis often use bet-based choice tasks without necessarily
853 incentivising them (e.g. Clemen and Reilly, 2013); again, our method, applied in this context,
854 complements existing stated approaches to eliciting probability intervals.

855 6. Conclusion

856 This paper proposes and implements a solution to the open problem of choice-based incentive-
857 compatible elicitation of multiple prior beliefs. It comprises a new preference-based notion—

858 Matching Probability Intervals—and probability-interval analogues of standard choice lists and
859 bisection elicitation procedures. Theoretically, it operates in the context of the Hurwicz α -
860 maxmin EU model and in the absence of strong assumptions about subjects' sets of priors, most
861 notably any form of probabilistic sophistication.

862 Our implementation of the elicitation method, in two experiments to elicit subjective upper
863 and lower CDFs over continuous-valued sources of uncertainty, testifies to its feasibility. It
864 finds a predominance of imprecision—a gap between upper and lower probabilities—across our
865 subjects, for all explored sources, showing it to be related to familiarity or predictability. It also
866 allows us to perform what, to our knowledge, is the first elicitation of the mixture coefficient in
867 the α -maxmin EU model that fully controls for beliefs.

868 A. Theoretical Appendix

869 In order to bring out the robustness of our proposal, and the assumptions underlying it, we shall
 870 at times work with a more general decision model than Hurwicz α -maxmin EU (1). Consider
 871 the representation where a bet $(z, E, 0)$ is evaluated according to:

$$W(\underline{p}(E), \bar{p}(E)).u(z) \tag{11}$$

872 where $\underline{p}(E)$, $\bar{p}(E)$ and u are as in Section 2.3, and W is an ‘aggregation function’, which is
 873 continuous in both coordinates and normalised— $W(x, x) = x$ for all x . α -maxmin EU is the
 874 special case where W is linear: $W(x, y) = \alpha x + (1 - \alpha)y$. See Grant et al. (2019) for an
 875 axiomatisation of a special case of (11) where the set of priors is generated by a probability
 876 measure on a subalgebra, and a thorough discussion of its potential. As in Section 2, we assume
 877 the same representation for imprecise risky prospects.

878 Note that, unlike α -maxmin EU, the general form (11) can accommodate non-linear, Prospect-
 879 Theory-style weighting of the lower and upper probabilities, for instance taking $W(x, y) =$
 880 $\alpha w(x) + (1 - \alpha)w(y)$, where w is a Prospect-Theory-style weighting function.

881 A.1. Proofs

882 We prove Proposition 1 under representation (11). As noted above, the α -maxmin EU model is
 883 a special case.

884 *Proof of Proposition 1.* Under (11), it follows from the first preference pattern in Proposition 1
 885 that $W(q, 1 - b) > W(\underline{p}(E), \bar{p}(E))$ for all $q > r$, and similarly for the others. By the continuity
 886 of W , it thus follows from the first two preferences that $W(r, 1 - b) = W(\underline{p}(E), \bar{p}(E))$, and
 887 from the second pair of preferences that $W(b, 1 - r) = W(1 - \bar{p}(E), 1 - \underline{p}(E))$. It thus follows
 888 that $(z, [r, 1 - b], 0) \sim (z, E, 0)$ and $(0, [r, 1 - b], z) \sim (0, E, z)$, so $[r, 1 - b]$ is a MPI for E , as
 889 required. \square

890 Note that the converse of Proposition 1 holds—i.e. for any MPI, the preference pattern in
 891 the Proposition holds—whenever W is strictly increasing in the first coordinate; this is the case
 892 for α -maxmin EU model with $\alpha > 0$.

893 *Proof of Proposition 2.* Part a. Plugging in the representations (1) and (2), any $[x, y]$ in the
 894 R-B region satisfies:

$$\begin{aligned} \alpha x + (1 - \alpha)y &\geq \alpha \underline{p}(E) + (1 - \alpha)\bar{p}(E) \\ \alpha(1 - y) + (1 - \alpha)(1 - x) &\geq \alpha(1 - \bar{p}(E)) + (1 - \alpha)(1 - \underline{p}(E)) \end{aligned}$$

895 By basic algebra (add α times the first inequality to $(1 - \alpha)$ times the second), one obtains
 896 $(\alpha^2 - (1 - \alpha)^2)x \geq (\alpha^2 - (1 - \alpha)^2)\underline{p}(E)$, whence it follows, since $\alpha > \frac{1}{2}$, that $x \geq \underline{p}(E)$. Similarly,

897 one obtains $((1 - \alpha)^2 - \alpha^2)y \geq ((1 - \alpha)^2 - \alpha^2)\bar{p}(E)$, whence, since $\alpha > \frac{1}{2}$, $y \leq \bar{p}(E)$. A similar
 898 argument establishes the result for points in W .

899 Part b. follows directly from the fact that, under 2, whenever $x \leq x'$ and $y \leq y'$, then
 900 $(z, [x, y], 0) \preceq (z, [x', y'], 0)$ and $(0, [x, y], z) \succeq (0, [x', y'], z)$. \square

901 We state for completeness the result on the uniqueness of the MPI.

902 **Proposition 3.** *For any decision maker represented according to (1) with $\alpha \neq \frac{1}{2}$, and for any*
 903 *event E , there is a unique MPI for E .*

904 *Proof.* Existence is immediate from Eqs. (3) and (4). Uniqueness is immediate from the linearity
 905 of the indifference curves in \mathcal{I} -space (see Figure 1). \square

906 A.2. Binary-choice procedure

907 A.2.1. Introduction and setup

908 Our binary-choice procedure is fully described in Figures 10–13. Figure 10 sets out the general
 909 structure (and stopping rules). At each step of the procedure, preferences are elicited for a single
 910 probability interval $[p_i, \bar{p}_i]$: i.e. preferences between the bet on the event and the IL $(z, [p_i, \bar{p}_i], 0)$,
 911 and between the bet on the complement event and the complementary IL $(0, [p_i, \bar{p}_i], z)$. The
 912 heart of the procedure, detailed in Figures 11–13, involves specification of the next probability
 913 interval proposed for elicitation on the basis of the preferences concerning the previous intervals.
 914 We first set out the notation used in the presentation of these parts of the procedure, before
 915 explaining informally its main steps. Throughout, we adopt the Euclidean topology on $\mathcal{I} \subseteq \mathbb{R}^2$,
 916 and let $d(\bullet, \bullet)$ be the Euclidean distance.

917 The procedure draws on two formal elements. The first is the assignment of interval-valued
 918 urns—or equivalently probability intervals—to preference-defined regions, discussed in Section
 919 2.6. Recall from Section 2.1 that an interval-valued urn $[p, q]$, i.e. with a minimum proportion p
 920 of red balls and a minimum proportion $1 - q$ of blue balls, corresponds to a probability interval;
 921 we shall present the procedure in terms of the latter here. For every event E_i and urn $[p, q]$,
 922 the preferences in the choices between the bet on E_i and that on a red ball being drawn from
 923 the urn, and between the bet on E_i^c and that on blue from the urn suffice to situate $[p, q]$ in
 924 one of the four regions, $R - B, W, R, B$ defined in Table 1 (Section 2.6). For instance, in Figure
 925 9, which we shall use to illustrate the procedure, the probability intervals already elicited are
 926 the dots coloured white, red, blue and red-blue according to the (preference-based) region they
 927 belong to.

928 The second element is a ‘polar’-style coordinate system for the set of probability intervals
 929 \mathcal{I} , under which, informally, $(m, \alpha) \in [0, 0.5] \times [0, 1]$ is the probability interval that is α along
 930 the piecewise-linear line that goes through the probability intervals $[0, 0]$, $[1, 1]$, and $[m, 1 - m]$
 931 (corresponding to the urn with at least proportion m of red balls and at least proportion m of
 932 blue balls). The thick grey line in Figure 9 is one such line. Formally, $\sigma : \mathcal{I} \rightarrow [0, 0.5] \times [0, 1]$ is
 933 defined by:

$$\sigma([p, q]) = \begin{cases} \left(\frac{p}{p+q}, \frac{p+q}{2}\right) & p \leq 1 - q, p + q \in (0, 2) \\ \left(\frac{1-q}{2-p-q}, \frac{p+q}{2}\right) & p > 1 - q, p + q \in (0, 2) \\ (0, 0) & p = q = 0 \\ (0, 1) & p = q = 1 \end{cases} \quad (12)$$

934 It is straightforward to check that σ is a well-defined function on \mathcal{I} . Every point except for
 935 $[0, 0], [1, 1]$ corresponds to a unique line (parametrised by m) and ‘distance’ along that line
 936 (parametrised by α). $[0, 0]$ (respectively $[1, 1]$) corresponds to a single α , namely 0 (resp. 1),
 937 though it lies on all such lines; we set the corresponding $m = 0$ by convention. For information,
 938 the inverse map is given by:

$$\sigma^{-1}(m, \alpha) = \begin{cases} [2\alpha m, 2\alpha(1 - m)] & \alpha \leq \frac{1}{2} \\ [(2 - 2\alpha)m + (2\alpha - 1), (2 - 2\alpha)(1 - m) + (2\alpha - 1)] & \alpha > \frac{1}{2} \end{cases} \quad (13)$$

939 We write $\sigma_1([p, q])$ (respectively $\sigma_2([p, q])$) for the first (resp. second coordinate) of $\sigma([p, q])$.
 940 Since this is a simple change of coordinates, we shall write $(m, \alpha) \in B$ as short for $\sigma^{-1}(m, \alpha) \in B$,
 941 and similarly for other cases.

942 A.2.2. Presentation of main steps

943 As discussed in Section 2.6 (Proposition 2), elicited points in the R-B and W regions determine
 944 an area in \mathcal{I} ‘between the R-B and the W points’ to which the MPI must belong. The general aim
 945 of the procedure is thus to find progressively ‘closer’ points in R-B and W, hence reducing the
 946 size of this area. This motivates the two main steps in the determination of the next probability
 947 interval to be presented for elicitation, $[p_{i+1}, q_{i+1}]$, on the basis of the previously elicited point
 948 $[p_i, q_i]$.

949 On the one hand, if $[p_i, q_i]$ is in the R-B region (respectively, the W region), then by Propo-
 950 sition 2 a. (Section 2.6), the MPI will be North-West of $[p_i, q_i]$ (resp. South-East of $[p_i, q_i]$) in
 951 Figure 1—i.e. $\underline{p} \leq p_i$ and $\bar{p} \geq q_i$ (resp. $\underline{p} \geq p_i$ and $\bar{p} \leq q_i$), where the MPI is $[\underline{p}, \bar{p}]$. In such cases,
 952 the procedure proposes a $[p_{i+1}, q_{i+1}]$ North-West (resp. South-East) of $[p_i, q_i]$. This exemplified
 953 by the $[p_{i+1}, q_{i+1}]$ proposed for point X in Figure 9. The precise proposal for $[p_{i+1}, q_{i+1}]$ depends
 954 on whether there is a point in W (resp. R-B); technicalities aside, this is the general strategy
 955 of the cases in lines 20-23 and 36-39 of the procedure (Figures 12-13). If the point $[p_{i+1}, q_{i+1}]$
 956 turns out to be in R-B or W, this will further restrict the area where the MPI can lie.

957 On the other hand, if $[p_i, q_i]$ is in the R or B regions, then Proposition 2 a. does not apply;
 958 as discussed in Section 2.6, the aim in such cases is to find a point in the R-B or W regions, to
 959 continue reducing the area containing the MPI. The procedure draws on two observations. First,
 960 as mentioned above, any point $[p_i, q_i]$ can be equivalently written in another coordinate system,
 961 specifying the line it sits on—parametrised by $m = \sigma_1([p_i, q_i])$ —and how ‘far’ along the line it

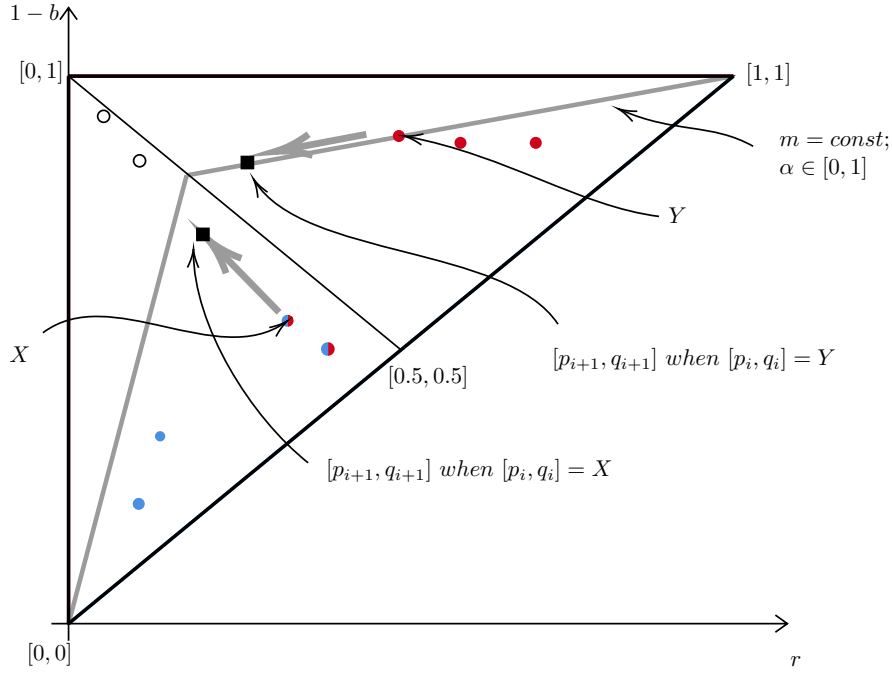


Figure 9: Binary Choice Procedure.

962 is—parametrised by $\alpha = \sigma_2([p_i, q_i])$. Second, for $[p_i, q_i]$ in R (respectively B), by Proposition
 963 2 b., all points North-East (resp. South-West) of $[p_i, q_i]$ are also in R (resp. B). So the only
 964 points in R-B and W on the line $m = \sigma_1([p_i, q_i])$ corresponding to the point $[p_i, q_i]$ must be
 965 South-West of $[p_i, q_i]$, i.e. with lower α (resp. North-East, i.e. with higher α). Accordingly, the
 966 procedure proposes a point $[p_{i+1}, q_{i+1}]$ on the line $m = \sigma_1([p_i, q_i])$ but shifted in the appropriate
 967 direction, as illustrated by the $[p_{i+1}, q_{i+1}]$ proposed for point Y (lying in the R region) in Figure
 968 9. Technicalities aside, this is general strategy for Case 1 (lines 1-17) and the cases in lines 24-34
 969 and lines 40-44 of the procedure (Figures 11–13). Among these cases, all retain the same m
 970 (grey line in Figure 9) except those considered in lines 12-17. These treat cases where no point
 971 in R-B or W has yet been found; the procedure in these cases increases m during the search,
 972 hence looking closer to the 45° line (ie. the line of $[p, q]$ with $p = q$). We used a procedure with
 973 this in-built precision bias to favour Bayesian replies (i.e. precise probabilities); in the light of
 974 it, our finding of widespread imprecision (Section 4.3) is all the more remarkable.

975 A.2.3. Convergence

976 Except for extreme cases, the procedure tends to the MPI.

977 **Proposition 4.** *Suppose preferences are represented according to (1) with $1 > \alpha > \frac{1}{2}$, let E*
 978 *be an event, and let $[\underline{p}_n, \bar{p}_n]$ be the result of the procedure in Figures 11–13 (with initial values*
 979 *set as in Figure 10) applied for n steps. Then $[\underline{p}_n, \bar{p}_n] \rightarrow [\underline{p}(E), \bar{p}(E)]$ as $n \rightarrow \infty$. Moreover,*
 980 *the procedure also converges in this sense when preferences are represented according to (1) with*
 981 *$\alpha = 1$, $\underline{p}(E) \neq 0$ and $\bar{p}(E) \neq 1$.*

982 *Proof.* We provide the main steps of the proof here; they rely on technical Lemmas 1–4, which are

983 detailed in Appendix D. We adopt the notation and initial values from Figure 10; in particular,
 984 let El_n be the set of elicited points after n steps. As discussed in Section 2.4, the MPI is
 985 $[\underline{p}(E), \bar{p}(E)]$. Moreover, by Proposition 2, at stage n , the MPI is contained in

$$\Phi_n = \left\{ [p, q] \in \mathcal{I} : \begin{array}{l} \max \{p' : [p', q'] \in El_n \cap W\} \leq p \leq \min \{p'' : [p'', q''] \in El_n \cap R - B\}, \\ \max \{q'' : [p'', q''] \in El_n \cap R - B\} \leq q \leq \min \{q' : [p', q'] \in El_n \cap W\} \end{array} \right\} \quad (14)$$

986 where the maximum of an empty set is taken to be 0 and the minimum 1.

987 We reason referring to the cases in the procedure (Figures 11–13). At the beginning of the
 988 procedure, it is in Case 1 ($El_0 \cap W = El_0 \cap R - B = \emptyset$). By lines 13-16, if no point in W or R-B
 989 is found, the points elicited by the procedure will reach the space of precise probabilities (i.e.
 990 points $[p, q]$ with $p = q$), where it will follow a standard bisection procedure. All such points
 991 have σ_1 -value of 0.5. It follows from Lemma 1 that if the MPI is not precise, then a point will be
 992 found in R-B, so the procedure moves to Case 2. On the other hand, if the MPI is precise, then,
 993 by Lemma 1 and the bisection character of the procedure on the space of precise probabilities,
 994 the points elicited in the procedure will converge to it as required.

995 Now consider cases where the procedure arrives to Case 2 or 3, i.e. it finds a point in R-B
 996 or W. By Lemma 3, $\sigma_1([p_n, q_n]) \rightarrow \sigma_1([\underline{p}(E), \bar{p}(E)])$ as $n \rightarrow \infty$. We distinguish three cases.

- 997 • $\sigma_1([\underline{p}(E), \bar{p}(E)]) > 0$ and $\sigma_1([p_n, q_n]) \neq \sigma_1([\underline{p}(E), \bar{p}(E)])$ for all n . By Proposition 2 and
 998 the definition of σ (and in particular the slopes of the lines $\sigma_1([p, q]) = m$ for $m > 0$),
 999 it follows straightforwardly that $\min_{[p, q] \in El_n} d([\underline{p}(E), \bar{p}(E)], [p, q]) \rightarrow 0$ as $n \rightarrow \infty$, whence
 1000 $[\underline{p}_n, \bar{p}_n] \rightarrow [\underline{p}(E), \bar{p}(E)]$ as required.
- 1001 • $\sigma_1([\underline{p}(E), \bar{p}(E)]) > 0$ and $\sigma_1([p_i, q_i]) = \sigma_1([\underline{p}(E), \bar{p}(E)])$ for some i . By Lemma 1 and Case
 1002 2 (lines 24-33) and Case 3 (lines 40-43), the procedure will, from i onwards, only pass
 1003 through points with same σ_1 -value $\sigma_1([\underline{p}(E), \bar{p}(E)])$, where it will only find points in R and
 1004 B. Moreover, it follows a bisection-style procedure on the line $\sigma_1([p, q]) = \sigma_1([\underline{p}(E), \bar{p}(E)])$.
 1005 It follows from standard arguments, Lemma 1 and representation (1) that this procedure
 1006 converges to $[\underline{p}(E), \bar{p}(E)]$ as required.
- 1007 • $\sigma_1([\underline{p}(E), \bar{p}(E)]) = 0$ and $\alpha < 1$ in the representation (1). Suppose $\underline{p}(E) = 0$; the other
 1008 case ($\underline{p}(E) \neq 0$ and so $\bar{p}(E) = 1$) is treated similarly. By Lemma 1, $[p_n, q_n]$ contains a
 1009 subsequence of points in R-B, with σ_1 -value tending to 0. Since $\alpha < 1$, by representation
 1010 (1), for every $q < \bar{p}(E)$, there exists $p > 0$ such that $(z, [p, q], 0) \prec (z, E, 0)$, and hence such
 1011 that $[p, q]$ is not in R-B. Moreover, by the representation and Lower Stochastic Dominance,
 1012 for every $q > \bar{p}(E)$ and p , $(0, [p, q], z) \prec (0, [\underline{p}(E), \bar{p}(E)], z) \sim (0, E, z)$, so such $[p, q]$ are
 1013 not in R-B. It follows that the subsequence of $[p_n, q_n]$ consisting of points in R-B converges
 1014 to $[\underline{p}(E), \bar{p}(E)]$, so $[p_n, q_n] \rightarrow [\underline{p}(E), \bar{p}(E)]$ as required.

1015 □

Procedure Binary Choice Procedure: structure and stopping rules

```

1 Set  $[p_1, q_1] = [0.3, 0.7]$ ,  $i = 1$  and  $El = \emptyset$ ; /*  $[p_i, q_i]$  is the last
   interval for which preferences were elicited, and  $El \not\ni [p_i, q_i]$ 
   is the set of intervals for which preferences have been
   elicited previously. So  $El \cup \{[p_i, q_i]\}$  is the set of all
   elicited intervals, including the one under consideration.
   */
2 Set  $rb_{El} = \arg \max_{[p,q] \in El \cap RB} q$ ,  $w_{El} = \arg \min_{[p,q] \in El \cap W} q$ ,
    $r_{El} = \arg \min_{[p,q] \in El \cap R} q$ ,  $b_{El} = \arg \max_{[p,q] \in El \cap B} q$ ; /*  $rb_{El}$  is the
   highest elicited point in R-B,  $w_{El}$  is the lowest elicited
   point in W, etc. */
3 while  $|El| < 12$  do
4   while  $d(rb_{El}, w_{El}) \geq 0.15$  do
5     Elicit preferences for  $[p_i, q_i]$ ;
6     Execute algorithm 1;
7     Add  $[p_i, q_i]$  to  $El$ ;
8     if  $d(rb_{El}, w_{El}) < 0.15$  then
9       return  $[p_i, q_i]$ ;
10    Stop
11 if  $El \cap W = \emptyset$  and  $El \cap R - B = \emptyset$  then
12   return  $\frac{b_{El} + r_{El}}{2}$ 
13 if  $El \cap R - B \neq \emptyset$  and  $El \cap W = \emptyset$  then
14   return  $rb_{El}$ 
15 if  $El \cap W \neq \emptyset$  and  $El \cap R - B = \emptyset$  then
16   return  $w_{El}$ 
17 if  $El \cap R - B \neq \emptyset$  and  $El \cap W \neq \emptyset$  then
18   return  $\frac{rb_{El} + w_{El}}{2}$ 

```

Figure 10: Binary choice procedure: structure

Algorithm 1: Determination of Next Binary Choice

```

1 Case 1
2   if  $(El \cup \{[p_i, q_i]\}) \cap R - B = (El \cup \{[p_i, q_i]\}) \cap W = \emptyset$  then
3     if there is no  $[p, q] \in El$  with  $\sigma_1([p, q]) = m$ , where  $m = \sigma_1([p_i, q_i])$  then
4       if  $[p_i, q_i] \in R$  then
5          $[p_{i+1}, q_{i+1}] = \sigma^{-1}(m, \max\{\alpha_0 | (m_0, \alpha_0) \in El \cap B\});$ 
6       if  $[p_i, q_i] \in B$  then
7          $[p_{i+1}, q_{i+1}] = \sigma^{-1}(m, \min\{\alpha_0 | (m_0, \alpha_0) \in El \cap R\});$ 
8     if there exists  $[p, q] \in El$  with  $\sigma_1([p, q]) = m$ , where  $m = \sigma_1([p_i, q_i])$ , but there
9       exists no  $[p, q] \in El$  with  $[p, q] \in B$  and  $\sigma_1([p, q]) = m$  then
10         $[p_{i+1}, q_{i+1}] = \sigma^{-1}(m, \frac{\sigma_2([p_i, q_i])}{2});$ 
11     if there exists  $[p, q] \in El$  with  $\sigma_1([p, q]) = m$ , where  $m = \sigma_1([p_i, q_i])$ , but there
12       exists no  $[p, q] \in El$  with  $[p, q] \in R$  and  $\sigma_1([p, q]) = m$  then
13         $[p_{i+1}, q_{i+1}] = \sigma^{-1}(m, \frac{1 + \sigma_2([p_i, q_i])}{2});$ 
14     if there exists  $[p, q], [p', q'] \in El$  with  $[p, q] \in B, [p', q'] \in R$  and
15        $\sigma_1([p, q]) = \sigma_1([p', q']) = \sigma_1([p_i, q_i]) = m$  then
16         if  $[p_i, q_i] \in R$  then
17            $[p_{i+1}, q_{i+1}] = \begin{cases} ((m + 0.5)/2, (\sigma_2([p_i, q_i]) + \alpha')/2) & m \leq 0.44 \\ (0.5, (\sigma_2([p_i, q_i]) + \alpha')/2) & m > 0.44 \end{cases}$ 
           where  $\alpha' = \max\{\alpha_0 | (m, \alpha_0) \in El \cap B\};$ 
           if  $[p_i, q_i] \in B$  then
            $[p_{i+1}, q_{i+1}] = \begin{cases} ((m + 0.5)/2, (\sigma_2([p_i, q_i]) + \alpha'')/2) & m \leq 0.44 \\ (0.5, (\sigma_2([p_i, q_i]) + \alpha'')/2) & m > 0.44 \end{cases}$ 
           where  $\alpha'' = \min\{\alpha_0 | (m, \alpha_0) \in R\};$ 
17   end

```

Figure 11: Determination of Next Binary Choice: Part 1

Notation: σ defined in (12) and (13).

18 **Case 2**

19 **if either** $(El \cup \{[p_i, q_i]\}) \cap R - B \neq \emptyset$ **or** $El \cup \{[p_i, q_i]\} \cap W \neq \emptyset$, **but not both then**

20 **if** $[p_i, q_i] \in R - B$ **then**

21

$$[p_{i+1}, q_{i+1}] = \begin{cases} \left[\frac{p_i + (p_i + q_i - 1)}{2}, \frac{q_i + 1}{2} \right] & p_i + q_i > 1 \\ \left[\frac{p_i}{2}, \frac{q_i + (p_i + q_i)}{2} \right] & p_i + q_i \leq 1 \end{cases}$$

22 **if** $[p_i, q_i] \in W$ **then**

23

$$[p_{i+1}, q_{i+1}] = \left[\frac{2p_i + q_i}{2}, \frac{p_i + 2q_i}{2} \right]$$

24 **if** $[p_i, q_i] \in R$ **then**

25

if $El \cap R - B \neq \emptyset$ **then**

26

$$[p_{i+1}, q_{i+1}] = \sigma^{-1} \left(\sigma_1([p_i, q_i]), \frac{\sigma_2([p_i, q_i]) + \alpha'}{2} \right)$$

where

$$\alpha' = \begin{cases} \frac{q'}{2(1-\sigma_1([p_i, q_i]))} & q' \leq 1 - \sigma_1([p_i, q_i]) \\ \frac{q'-1}{2\sigma_1([p_i, q_i])} + 1 & q' > 1 - \sigma_1([p_i, q_i]) \end{cases}$$

with $[p', q'] = \arg \max_{[p, q] \in El \cap R - B} q$.

27

if $El \cap W \neq \emptyset$ **then**

28

$$[p_{i+1}, q_{i+1}] = \sigma^{-1} \left(\sigma_1([p_i, q_i]), \frac{\sigma_2([p_i, q_i]) + \alpha''}{2} \right)$$

where

$$\alpha'' = \begin{cases} \frac{p''}{2\sigma_1([p_i, q_i])} & p'' \leq \sigma_1([p_i, q_i]) \\ \frac{p'' + 1 - 2\sigma_1([p_i, q_i])}{2(1-\sigma_1([p_i, q_i]))} & p'' > \sigma_1([p_i, q_i]) \end{cases}$$

with $[p'', q''] = \arg \max_{[p, q] \in El \cap W} p$.

29

if $[p_i, q_i] \in B$ **then**

30

if $El \cap R - B \neq \emptyset$ **then**

31

$$[p_{i+1}, q_{i+1}] = \sigma^{-1} \left(\sigma_1([p_i, q_i]), \frac{\sigma_2([p_i, q_i]) + \alpha''}{2} \right)$$

where

$$\alpha'' = \begin{cases} \frac{p''}{2\sigma_1([p_i, q_i])} & p'' \leq \sigma_1([p_i, q_i]) \\ \frac{p'' + 1 - 2\sigma_1([p_i, q_i])}{2(1-\sigma_1([p_i, q_i]))} & p'' > \sigma_1([p_i, q_i]) \end{cases}$$

with $[p'', q''] = \arg \min_{[p, q] \in El \cap R - B} p$.

32

if $El \cap W \neq \emptyset$ **then**

33

$$[p_{i+1}, q_{i+1}] = \sigma^{-1} \left(\sigma_1([p_i, q_i]), \frac{\sigma_2([p_i, q_i]) + \alpha'}{2} \right)$$

where

$$\alpha' = \begin{cases} \frac{q'}{2(1-\sigma_1([p_i, q_i]))} & q' \leq 1 - \sigma_1([p_i, q_i]) \\ \frac{q'-1}{2\sigma_1([p_i, q_i])} + 1 & q' > 1 - \sigma_1([p_i, q_i]) \end{cases}$$

with $[p', q'] = \arg \min_{[p, q] \in El \cap W} q$.

34

end

Figure 12: Determination of Next Binary Choice: Part 2

34 Case 3
35 if $(EI \cup \{[p_i, q_i]\}) \cap R - B \neq \emptyset$ **and** $(EI \cup \{[p_i, q_i]\}) \cap W \neq \emptyset$ **then**
36 if $[p_i, q_i] \in R - B$ **then**
37

$$[p_{i+1}, q_{i+1}] = \left[\frac{p_i + p''}{2}, \frac{q_i + q''}{2} \right]$$

with $[p'', q''] = \arg \max_{[p, q] \in EI \cap W} p$.

38 if $[p_i, q_i] \in W$ **then**
39

$$[p_{i+1}, q_{i+1}] = \left[\frac{p_i + p'}{2}, \frac{q_i + q'}{2} \right]$$

with $[p', q'] = \arg \min_{[p, q] \in EI \cap RB} p$.

40 if $[p_i, q_i] \in B$ **then**
41

$$[p_{i+1}, q_{i+1}] = \sigma^{-1} \left(\sigma_1([p_i, q_i]), \frac{\sigma_2([p_i, q_i]) + \max(\alpha', \alpha'')}{2} \right)$$

where

$$\alpha' = \begin{cases} \frac{q'}{2(1-\sigma_1([p_i, q_i]))} & q' \leq 1 - \sigma_1([p_i, q_i]) \\ \frac{q'-1}{2\sigma_1([p_i, q_i])} + 1 & q' > 1 - \sigma_1([p_i, q_i]) \end{cases}$$

with $[p', q'] = \arg \max_{[p, q] \in EI \cap R - B} q$ **and**

$$\alpha'' = \begin{cases} \frac{p''}{2\sigma_1([p_i, q_i])} & p'' \leq \sigma_1([p_i, q_i]) \\ \frac{p''+1-2\sigma_1([p_i, q_i])}{2(1-\sigma_1([p_i, q_i]))} & p'' > \sigma_1([p_i, q_i]) \end{cases}$$

with $[p'', q''] = \arg \min_{[p, q] \in EI \cap W} q$.

42 if $[p_i, q_i] \in R$ **then**
43

$$[p_{i+1}, q_{i+1}] = \sigma^{-1} \left(\sigma_1([p_i, q_i]), \frac{\sigma_2([p_i, q_i]) + \min(\alpha', \alpha'')}{2} \right)$$

where

$$\alpha' = \begin{cases} \frac{q''}{2(1-\sigma_1([p_i, q_i]))} & q'' \leq 1 - \sigma_1([p_i, q_i]) \\ \frac{q''-1}{2\sigma_1([p_i, q_i])} + 1 & q'' > 1 - \sigma_1([p_i, q_i]) \end{cases}$$

with $[p'', q''] = \arg \min_{[p, q] \in EI \cap W} q$ **and**

$$\alpha'' = \begin{cases} \frac{p'}{2\sigma_1([p_i, q_i])} & p' \leq \sigma_1([p_i, q_i]) \\ \frac{p'+1-2\sigma_1([p_i, q_i])}{2(1-\sigma_1([p_i, q_i]))} & p' > \sigma_1([p_i, q_i]) \end{cases}$$

with $[p', q'] = \arg \max_{[p, q] \in EI \cap R - B} q$.

44 end

Figure 13: Determination of Next Binary Choice: Part 3

1016 **A.3. Robustness of the elicitation method**

1017 As stated in Section 2, the proposed elicitation method has three novel elements. The first is
 1018 the notion of MPI, and the observation that they yield the probability intervals generated by
 1019 the subjects’ set of priors. The second is the incentivisation mechanism, based on the 2D choice
 1020 list set out in Section 2.5. As for elicitation of subjective probabilities (e.g. choice-list methods
 1021 for eliciting MPs), this is already sufficient to provide an elicitation mechanism for subjects’ sets
 1022 of priors. However, the proposal also includes a chained binary-choice procedure, in the style
 1023 of the ‘bisection’ or ‘staircase’ method for MPs or certainty equivalents, to aid the subject find
 1024 the MPI (Section 3.4). We now discuss to what extent the proposed elements apply beyond
 1025 the typical α -maxmin EU representation with $\alpha > \frac{1}{2}$ on which we have focused in Section 2.
 1026 Whilst we concentrate below on extensions to models of the more general form (11), note that
 1027 the method also applies under other multiple-prior decision models, most notably multiple-prior
 1028 minimax (expected) regret.¹⁶ We also further analyse the incentive compatibility properties of
 1029 the 2D choice list incentivisation mechanism.

1030 **A.3.1. Matching Probability Intervals**

1031 Under the general preferences of the form (11), the equations (5) for the MPI can be rewritten
 1032 in the obvious way.¹⁷ Clearly, the notion of MPI is well defined, and the subjective probability
 1033 interval is an MPI. The form of W can however affect the uniqueness of the MPI. More precisely,
 1034 it is guaranteed to be unique whenever there is a unique solution to the equations, and this
 1035 only occurs if W satisfies the following ‘single-crossing property’: every pair of red-and-blue
 1036 indifference curves in Figure 1 cross at most once.¹⁸ Whether this is the case, and how often
 1037 it is not, will depend on the functional form of W . We thus consider what form of uniqueness
 1038 holds for reasonable W .

1039 For instance, the MPI is clearly unique when W is linear and non-symmetric¹⁹—and hence
 1040 for α -maxmin EU whenever $\alpha \neq \frac{1}{2}$. A more general interesting case is when W incorporates
 1041 probability weighting, e.g. is of the form $W(x, y) = \alpha w(x) + (1 - \alpha)w(y)$ for a weighting

¹⁶This model evaluates the choice of act f from a menu M according to $-\max_{p \in \mathcal{C}} \mathbb{E}_p(\max_{g \in M} u(g(s)) - u(f(s)))$, where \mathbb{E}_p is the expectation with respect to probability measure p and \mathcal{C} is the set of priors (e.g. Berger, 1985; Stoye, 2011). It is straightforward to show that for the choices used by our method—namely binary choices between bets on independent events, in the sense that the joint (multi-prior) distribution over the pair of relevant events is a ‘type-1 product’ (Walley, 1991, Sect. 9.3.5) of the multiple priors beliefs about each—preferences under this rule correspond to preferences under maxmin-EU (i.e. (1) with $\alpha = 1$) with the same set of priors.

¹⁷Explicitly:

$$W(\underline{p}, \bar{p}) = W(\underline{p}(E), \bar{p}(E)) \tag{15}$$

$$W(1 - \bar{p}, 1 - \underline{p}) = W(1 - \bar{p}(E), 1 - \underline{p}(E)) \tag{16}$$

¹⁸Technically, for every $A, B \in \mathbb{R}$, $|\{[x, y] \in \mathcal{I} : W(x, y) = A, W(1 - y, 1 - x) = B\}| \leq 1$.

¹⁹I.e. it is not the case that $W(x, y) = W(y, x)$ for all x, y .

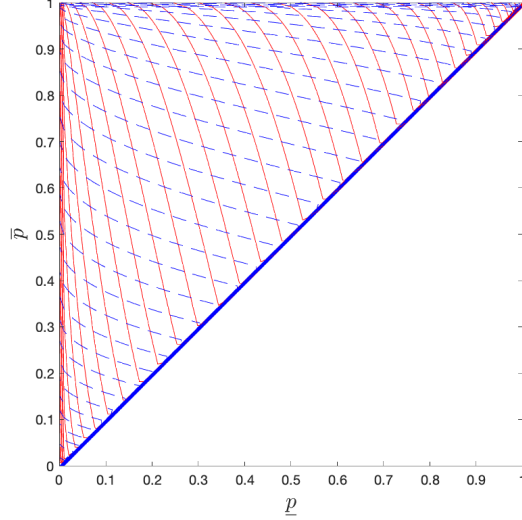


Figure 14: Indifference curves in probability interval space \mathcal{I} under (11) with $W(x, y) = \alpha w(x) + (1 - \alpha)w(y)$.

Red lines: indifference curves for IL $(z, [p, q], 0)$: i.e. curves of the form $\alpha w(x) + (1 - \alpha)w(y) = C$.

Blue lines: indifference curves for IL $(0, [p, q], z)$: i.e. curves of the form $\alpha w(1 - y) + (1 - \alpha)w(1 - x) = D$.

Parametrisation: Prelec weighting function $w(x) = (e^{-(-\ln(x))^\alpha})^\beta$ with $\alpha = 0.54$ and $\beta = 0.85$ (Abdellaoui et al., 2011); $\alpha = 0.8$.

1042 function w . Note that this form can incorporate findings on probability weighting for (two-
 1043 outcome) lotteries, via w . For such W , if w takes the quasi-linear form often used in literature
 1044 (Chateauneuf et al., 2007; Wakker, 2010), then MPIs can be shown to remain unique (by a
 1045 similar reasoning to that for the non-weighted case). Moreover, even for non-linear weighting
 1046 functions, calculation of relevant cases suggests that MPIs are typically unique. As an example,
 1047 Figure 14 plots red and blue indifference curves for the specified form of W with w being the
 1048 popular Prelec weighting function with the parameters found by Abdellaoui et al. (2011) for a
 1049 Paris temperature source—i.e. one that is similar to the source we used in EXP 1—and an α
 1050 of 0.8—i.e. close to the value we found for α (Section 4.5). Clearly, red and blue indifference
 1051 curves typically only cross (at most) once, as required for uniqueness of MPI. Even in the cases
 1052 where there are multiple MPIs, there will be at most two, with one close to the horizontal or
 1053 vertical boundary.

1054 In summary, even for reasonable extensions beyond α -maxmin EU, MPIs are well-defined,
 1055 and the subject's probability interval is always a MPI. Moreover, there is reason to believe that
 1056 uniqueness continues to hold largely, and where it does not, there is at most one other possible
 1057 candidate MPI. Note that even in cases of non-uniqueness, the analysis of the 2D choice list
 1058 incentivisation mechanism is unaffected, and every MPI remains a weakly dominant strategy.
 1059 So it will yield a candidate probability interval.

1060 **A.3.2. 2D Choice List**

1061 The 2D choice list set out in Section 2.5 is incentive compatible under the α -maxmin EU model
 1062 (1) whenever subjects treat the two branches of the choice list in isolation from each other.
 1063 We first consider the consequences of violation of this isolation assumption, before turning to
 1064 robustness to generalisation of the decision model.

1065 **Robustness to violations of isolation** Suppose that the isolation assumption discussed in
 1066 Section 5 does not hold, and the subject reasons strategically across the two branches of the 2D
 1067 choice list. Then the choice of MPI is conceptualised as the choice of a (second-order) lottery
 1068 assigning a probability to playing a bet for or against E or to playing specific ILs according to
 1069 the mechanism. Assuming the α -maximin EU model (1) at both levels, the subject evaluates
 1070 each such second-order lottery using the expectation over the values of the bets and ILs. Let
 1071 $[p(E), \bar{p}(E)] = [\underline{p}, \bar{p}]$. For any reported point $[q, \bar{q}]$ in this task, by the incentive mechanism
 1072 defined in Section 2.5:

- 1073 • the probability of receiving the bet on E is $\frac{q}{\bar{q}+1-q}$
- 1074 • the probability of receiving the IL on red is $\frac{\bar{q}-q}{\bar{q}+1-q}$
- 1075 • the probability of receiving the bet on E^c is $\frac{1-\bar{q}}{\bar{q}+1-q}$
- 1076 • the probability of receiving the IL on red is $\frac{\bar{q}-q}{\bar{q}+1-q}$

1077 Using these and the evaluations of the bets and the ILs according to (1) and (2) (with $[\underline{p}, \bar{p}]$),
 1078 one obtains the following form for the utility of reporting $[q, \bar{q}]$ when the true beliefs are $[\underline{p}, \bar{p}]$:

$$\begin{aligned} & \frac{(1-\bar{q})(\alpha(1-\bar{p})-(\alpha-1)(1-\underline{p}))}{\bar{q}-\underline{q}+1} + \frac{(\bar{q}-q)\left(\frac{q}{2} + \frac{\alpha(1-\bar{q})}{2} - \frac{(\alpha-1)(1-q)}{2} - \frac{1}{2}\right)}{\bar{q}-q+1} \\ & + \frac{q(\alpha\underline{p}-\bar{p}(\alpha-1))}{\bar{q}-q+1} + \frac{(\bar{q}-q)\left(\frac{\bar{q}}{2} + \frac{\alpha q}{2} - \frac{\bar{q}(\alpha-1)}{2}\right)}{\bar{q}-q+1} \end{aligned}$$

1079 Finding the optimum numerically for a grid of values of $\underline{p}, \bar{p}, \alpha \in [0, 1]$ using Matlab, we find that,
 1080 for every $(\underline{p}, \bar{p}, \alpha)$ (with $\bar{p} \geq \underline{p}$) except for $\underline{p} = 0, \bar{p} = 1, \alpha = 0$, and those with $\underline{p} = 0.5, \alpha = 1$
 1081 or $\bar{p} = 0.5, \alpha = 1$, the maximum is attained at one or several of the ‘vertices’ of the triangle in
 1082 Figure 1, i.e. $[0, 0], [0, 1], [1, 1]$. For $\underline{p} = 0, \bar{p} = 1, \alpha = 0$ and $\underline{p} = 0.5, \alpha = 1$ or $\bar{p} = 0.5, \alpha = 1$
 1083 with $\underline{p} \neq \bar{p}$, the maximum is attained at all points on one of the boundaries of the triangle, i.e.
 1084 $\{[0, y] : y \in [0, 1]\}, \{[x, 1] : x \in [0, 1]\}, \{[x, y] : x \in [0, 1], y = x\}$. When $\underline{p} = 0.5, \bar{p} = 0.5, \alpha = 1$,
 1085 the utility above is constant, so all points maximise it.

1086 It follows that, for any subject with $\alpha \in (0, 1)$ who violates the isolation assumption and
 1087 responds to the choice list strategically, every response will be at a vertex of the space \mathcal{I} . Our
 1088 elicitation of α suggests that the vast majority of subjects have α in this range. Even for subjects

1089 with $\alpha = 0$ or 1 reasoning strategically, they will have more than one response in the interior
1090 of \mathcal{I} if $\alpha = 1$ and they assign precise probability of 0.5 to several elicited events—which, in
1091 our experiment involving nested events, would correspond to a peculiar (bimodal) distribution
1092 across the variable of interest (temperature, marks). As is clear from Table 9 (Appendix B.1),
1093 no subjects give vertex responses for all elicited events, with only one subject (across both
1094 experiments) giving vertex responses for over half of the elicited events. Moreover, the vast
1095 majority of subjects (73 out of 80 in EXP 1; 51 out of 52 in EXP 2) gave more than one
1096 response in the interior of \mathcal{I} . The data thus clearly suggests that strategic reasoning is very
1097 limited in our sample.

1098 **Robustness to generalizations of the decision model** As suggested in Section 2.5, the
1099 incentivization mechanism implemented by the 2D choice list relies solely on the weak Lower
1100 Stochastic Dominance property (Definition 1). Formulated in terms of Eq. (11), this is just
1101 the assumption that W is strictly increasing in the first coordinate—or, in terms of preferences,
1102 decision makers are sensitive to the lower winning probability. The only reasonable model in
1103 the family of form (11) violating this property is the maxmax EU— α -maxmin EU with $\alpha = 0$.
1104 Since there is basically no evidence for a significant number of subjects with such preferences,
1105 the incentive compatibility of the 2D choice list discussed in Section 2.5 generalizes widely.

1106 A.3.3. Binary-choice procedure

1107 The binary-choice procedure (Section 2.6 and Appendix A.2) is based on the division of \mathcal{I} into
1108 regions, displayed in Table 1, and Proposition 2, in particular part a. dictating ‘where’ the
1109 MPI is relative to points in two of the regions (the W and R-B regions). For decision makers
1110 represented according to the α -maxmin EU model (1), Proposition 2 a. only holds if $\alpha > \frac{1}{2}$.²⁰
1111 When $\alpha < \frac{1}{2}$, the opposite of the statement in the Proposition holds: the MPI is North-West
1112 of the elicited point (on Figure 1) not when the latter is in R-B, but when it is in W (and
1113 similarly for South-East). So the algorithm applied to such decision makers would ‘move’ in
1114 the wrong direction: instead of looking South-East for the MPI after finding a point in R-B, it
1115 would look North-West, for instance. Note that, even if the algorithm does not work properly
1116 for such decision makers, the 2D choice list incentivisation mechanism is still valid, and hence
1117 they would, in principle, correct any issues at the 2D choice list confirmation stage. To gain
1118 some insight into the extent of procedure malfunction due to $\alpha < \frac{1}{2}$, we can look at the evidence
1119 on the value of α for our subjects, as well as some statistics on the functioning of the procedure.

1120 We find little evidence for widespread $\alpha < \frac{1}{2}$ among our subjects. First of all, the elicitation
1121 of α reported in Section 4.5 finds median and 25 percentile values significantly above $\frac{1}{2}$ (Table
1122 24), indicating that less than 25% of subjects have $\alpha < \frac{1}{2}$. Moreover, under the α -maxmin EU
1123 model, the sum of the MP of an event and that of its complement is less than (respectively,
1124 greater than) one precisely when $\alpha > \frac{1}{2}$ (resp. $\alpha < \frac{1}{2}$; see Appendix B.4), indicating that we

²⁰It also holds under the probability weighting specification of (11) mentioned in Section A.3.1 when $\alpha > \frac{1}{2}$.

1125 can use our matching probability data to check for the sign of $\alpha - \frac{1}{2}$. Table 25 (Appendix B.4)
1126 displays the descriptive statistics on this sum for the Paris treatment where MPs were elicited,
1127 confirming again that $\alpha > \frac{1}{2}$ for over 75% of subjects.

1128 As concerns its functioning, since the procedure ‘moves’ in the wrong direction for subjects
1129 with $\alpha < \frac{1}{2}$, no such subjects will pass through both points in W and points in R-B during
1130 the procedure. However, 383 applications of the procedure out of 704 in EXP 1 passed through
1131 points in W and R-B (300 out of 606 in EXP 2). Whilst there were nevertheless applications
1132 of the procedure which passed through points in R-B but not W (152 in EXP 1, 77 in EXP 2)
1133 and in W but not R-B (114 in EXP 1, 105 in EXP 2), these would be expected if the procedure
1134 functioned correctly and the probability intervals were large (respectively small). The evidence
1135 thus does not support a hypothesis involving malfunctioning of the procedure over explanations,
1136 such as this, relating to proper functioning and the character of the elicited intervals.

1137

1138 B. Supplementary Statistics

1139 B.1. Descriptive Statistics

1140 Tables 5–8 report the basic descriptive statistics on the upper and lower elicited probabilities
1141 after the ‘confirmation’ 2D choice list, and before the confirmation screen but after the binary-
1142 choice procedure, respectively.

1143 **Elicited points on a vertex** Table 9 reports counts of the number of subjects with a given
1144 number of elicited points at the vertex of the space \mathcal{I} of partially known urns (and corresponding
1145 probability intervals) in Figure 1.

1146 **Monotonicity** Tables 10a and 10b report the descriptive statistics for the individuallevel
1147 Kendall τ_b , calculated over the events in each source.

1148 B.1.1. Imprecision

1149 Table 11 presents the descriptive statistics for the Imprecision Index, whereas Table 12 displays
1150 counts of the number of subjects with various numbers of precise elicited points, as well as
1151 differences before the 2D choice list confirmation stage of the experiment as opposed to after.
1152 Table 13 presents the results of ANOVAs of the imprecision concerning an event against the
1153 event, for each source, where the null hypothesis is that imprecision is invariant across events.

1154

Event E_t for $t =$	count	mean	std	min	25%	50%	75%	max
-2	80.0	0.29	0.17	-0.01	0.15	0.29	0.40	0.7
2	80.0	0.38	0.22	0.00	0.20	0.35	0.50	1.0
5	80.0	0.48	0.23	0.00	0.35	0.46	0.66	1.0
8	80.0	0.57	0.24	0.05	0.42	0.59	0.75	1.0

(a) Lower probabilities Paris

Event E_t for $t =$	count	mean	std	min	25%	50%	75%	max
-2	80.0	0.55	0.21	0.09	0.40	0.55	0.67	1.0
2	80.0	0.65	0.19	0.23	0.51	0.65	0.80	1.0
5	80.0	0.74	0.17	0.25	0.62	0.76	0.88	1.0
8	80.0	0.82	0.14	0.50	0.75	0.85	0.94	1.0

(b) Upper probabilities Paris

Event E_t for $t =$	count	mean	std	min	25%	50%	75%	max
15	80.0	0.31	0.22	0.00	0.14	0.26	0.45	0.95
18	80.0	0.35	0.26	0.00	0.14	0.32	0.47	1.00
20	80.0	0.41	0.27	-0.01	0.20	0.40	0.61	1.00
22	80.0	0.43	0.26	-0.01	0.20	0.39	0.61	1.00

(c) Lower probabilities Sydney

Event E_t for $t =$	count	mean	std	min	25%	50%	75%	max
15	80.0	0.58	0.27	0.01	0.37	0.56	0.84	0.99
18	80.0	0.66	0.24	0.03	0.50	0.69	0.88	1.00
20	80.0	0.71	0.23	0.01	0.60	0.76	0.89	1.00
22	80.0	0.73	0.23	0.00	0.58	0.80	0.92	1.00

(d) Upper probabilities Sydney

Table 5: Descriptive Statistics: Elicited lower and upper probabilities after 2D choice list, EXP

1

Event E_t for $t =$	count	mean	std	min	25%	50%	75%	max
7	52.0	0.05	0.07	-0.01	0.00	0.04	0.08	0.35
10	52.0	0.15	0.12	0.00	0.06	0.12	0.19	0.50
12	52.0	0.24	0.15	0.00	0.14	0.20	0.31	0.63
15	52.0	0.40	0.19	0.08	0.26	0.38	0.55	0.73
17	52.0	0.60	0.16	0.18	0.54	0.64	0.71	0.86

(a) Lower probabilities Maths

Event E_t for $t =$	count	mean	std	min	25%	50%	75%	max
7	52.0	0.12	0.14	0.00	0.02	0.08	0.16	0.60
10	52.0	0.23	0.17	0.00	0.10	0.20	0.30	0.65
12	52.0	0.35	0.18	0.04	0.22	0.32	0.48	0.72
15	52.0	0.54	0.20	0.08	0.40	0.52	0.70	0.87
17	52.0	0.75	0.15	0.22	0.65	0.78	0.86	1.01

(b) Upper probabilities Maths

Event E_t for $t =$	count	mean	std	min	25%	50%	75%	max
7	52.0	0.12	0.08	0.01	0.06	0.11	0.17	0.36
10	52.0	0.22	0.12	0.02	0.14	0.20	0.29	0.50
12	52.0	0.33	0.13	0.14	0.22	0.32	0.40	0.60
15	52.0	0.54	0.14	0.19	0.46	0.56	0.65	0.83
17	52.0	0.71	0.13	0.25	0.65	0.74	0.83	0.90

(c) Lower probabilities Contraction

Event E_t for $t =$	count	mean	std	min	25%	50%	75%	max
7	52.0	0.22	0.14	0.02	0.11	0.20	0.31	0.56
10	52.0	0.37	0.14	0.06	0.30	0.34	0.46	0.65
12	52.0	0.51	0.14	0.20	0.40	0.50	0.64	0.74
15	52.0	0.74	0.11	0.40	0.67	0.77	0.82	0.90
17	52.0	0.86	0.07	0.60	0.82	0.86	0.91	1.00

(d) Upper probabilities Contraction

Table 6: Descriptive Statistics: Elicited lower and upper probabilities after 2D choice list, EXP

Event E_t for $t =$	count	mean	std	min	25%	50%	75%	max
-2	80.0	0.28	0.15	0.0	0.15	0.30	0.39	0.64
2	80.0	0.36	0.21	0.0	0.20	0.33	0.45	1.00
5	80.0	0.48	0.23	0.0	0.35	0.46	0.67	1.00
8	80.0	0.54	0.25	0.05	0.35	0.55	0.74	1.00

(a) Lower probabilities Paris

Event E_t for $t =$	count	mean	std	min	25%	50%	75%	max
-2	80.0	0.55	0.21	0.0	0.39	0.56	0.68	0.99
2	80.0	0.64	0.18	0.23	0.52	0.67	0.78	1.00
5	80.0	0.75	0.17	0.25	0.61	0.79	0.89	1.00
8	80.0	0.81	0.14	0.50	0.73	0.85	0.92	1.00

(b) Upper probabilities Paris

Event E_t for $t =$	count	mean	std	min	25%	50%	75%	max
15	80.0	0.28	0.21	0.0	0.12	0.26	0.38	0.95
18	80.0	0.33	0.27	0.0	0.11	0.30	0.45	1.00
20	80.0	0.43	0.28	0.0	0.19	0.42	0.61	1.00
22	80.0	0.42	0.26	0.0	0.21	0.35	0.61	1.00

(c) Lower probabilities Sydney

Event E_t for $t =$	count	mean	std	min	25%	50%	75%	max
15	80.0	0.60	0.27	0.00	0.42	0.61	0.83	0.99
18	80.0	0.67	0.25	0.03	0.50	0.66	0.89	1.00
20	80.0	0.73	0.23	0.00	0.60	0.79	0.89	1.00
22	80.0	0.74	0.22	0.00	0.62	0.80	0.91	1.00

(d) Upper probabilities Sydney

Table 7: Descriptive Statistics: Elicited lower and upper probabilities after binary-choice procedure and before 2D choice list, EXP 1

Event E_t for $t =$	count	mean	std	min	25%	50%	75%	max
7	52.0	0.11	0.16	0.00	0.01	0.06	0.14	1.00
10	52.0	0.18	0.14	0.00	0.08	0.14	0.30	0.50
12	52.0	0.25	0.15	0.00	0.15	0.22	0.32	0.60
15	52.0	0.38	0.22	0.04	0.19	0.33	0.53	1.00
17	52.0	0.54	0.18	0.07	0.43	0.60	0.66	0.86

(a) Lower probabilities Maths

Event E_t for $t =$	count	mean	std	min	25%	50%	75%	max
7	52.0	0.18	0.23	0.00	0.03	0.08	0.28	1.00
10	52.0	0.27	0.22	0.00	0.10	0.23	0.44	0.70
12	52.0	0.40	0.23	0.00	0.22	0.38	0.61	0.78
15	52.0	0.56	0.21	0.08	0.39	0.58	0.71	1.00
17	52.0	0.73	0.16	0.22	0.62	0.76	0.85	0.99

(b) Upper probabilities Maths

Event E_t for $t =$	count	mean	std	min	25%	50%	75%	max
7	52.0	0.15	0.09	0.01	0.08	0.15	0.20	0.40
10	52.0	0.26	0.12	0.00	0.18	0.23	0.35	0.48
12	52.0	0.32	0.16	0.04	0.21	0.30	0.40	0.71
15	52.0	0.53	0.19	0.16	0.40	0.55	0.66	1.00
17	52.0	0.67	0.20	0.07	0.62	0.68	0.83	1.00

(c) Lower probabilities Contraction

Event E_t for $t =$	count	mean	std	min	25%	50%	75%	max
7	52.0	0.30	0.23	0.03	0.12	0.24	0.41	0.85
10	52.0	0.41	0.18	0.00	0.31	0.35	0.56	0.86
12	52.0	0.54	0.16	0.20	0.44	0.55	0.65	0.87
15	52.0	0.74	0.14	0.37	0.66	0.74	0.84	1.00
17	52.0	0.85	0.09	0.50	0.82	0.86	0.91	1.00

(d) Upper probabilities Contraction

Table 8: Descriptive Statistics: Elicited lower and upper probabilities after binary-choice procedure and before 2D choice list, EXP 2

Point	Paris					Sydney				
	# subjects					# subjects				
	0	1	2	3	4	0	1	2	3	4
[0, 0]	80	0	0	0	0	80	0	0	0	0
[0, 1]	80	0	0	0	0	80	0	0	0	0
[1, 1]	79	0	0	1	0	78	1	0	1	0
[0, 0], [0, 1] or [1, 1]	79	0	0	1	0	78	1	0	1	0

(a) EXP 1

Point	Maths						Contraction					
	# subjects						# subjects					
	0	1	2	3	4	5	0	1	2	3	4	5
[0, 0]	46	5	1	0	0	0	52	0	0	0	0	0
[0, 1]	52	0	0	0	0	0	52	0	0	0	0	0
[1, 1]	52	0	0	0	0	0	52	0	0	0	0	0
[0, 0], [0, 1] or [1, 1]	46	5	1	0	0	0	52	0	0	0	0	0

(b) EXP 2

Table 9: For each type of point, the table indicates the number of subjects with the specified number of elicited points being of this type.

	MP Lower	MP Upper	Paris Lower	Paris Upper	Sydney Lower	Sydney Upper
count	74	78	79	78	78	78
mean	0.62	0.66	0.56	0.56	0.27	0.41
std	0.46	0.38	0.45	0.47	0.59	0.50
min	-0.91	-0.91	-0.91	-0.91	-1.00	-1.00
25%	0.55	0.55	0.33	0.33	-0.14	0.00
50%	0.71	0.69	0.67	0.67	0.33	0.55
75%	0.91	0.91	1.00	1.00	0.67	0.91
max	1.00	1.00	1.00	1.00	1.00	1.00

(a) EXP 1

	Contraction Lower	Contraction Upper	Maths Lower	Maths Upper
count	52	52	52	52
mean	0.99	0.99	0.98	1.00
std	0.02	0.03	0.07	0.01
min	0.95	0.80	0.53	0.95
25%	1.00	1.00	1.00	1.00
50%	1.00	1.00	1.00	1.00
75%	1.00	1.00	1.00	1.00
max	1.00	1.00	1.00	1.00

(b) EXP 2

Table 10: Individual-level Kendall τ_b descriptive statistics

Note τ_b is not defined for some subjects in EXP 1 (because of too many ties), and they were dropped.

	EXP 1		EXP 2	
	Paris	Sydney	Maths	Contraction
count	80	80	52	52
mean	0.25	0.29	0.13	0.19
std	0.17	0.20	0.09	0.11
min	0	0	0	0.01
25%	0.10	0.10	0.06	0.11
50%	0.23	0.28	0.11	0.17
75%	0.35	0.43	0.19	0.23
max	0.82	0.75	0.45	0.53

Table 11: Imprecision Index (Eq. (8)) descriptive statistics; EXP 1 and EXP 2.

# Precise events	EXP 1				EXP 2			
	Paris		Sydney		Maths		Contraction	
	2D C.L.	B-C Proc	2D C.L.	B-C Proc	2D C.L.	B-C Proc	2D C.L.	B-C Proc
0	51	48	48	44	20	12	31	19
1	14	14	18	23	14	14	12	20
2	7	11	8	8	12	10	5	3
3	6	0	2	4	3	10	3	4
4	2	1	4	1	0	2	1	7
5	-	-	-	-	3	4	0	2
Total	80	80	80	80	52	52	52	52

Table 12: Number of subjects with given number of precise events, per source. Data given after the 2D choice list confirmation screen (2D C.L.) and after the binary-choice procedure but before the confirmation screen (B-C Proc).

	Source	F	p-value
EXP 1	Paris	0.1048	0.957
	Sydney	0.4769	0.698
EXP 2	Contraction	4.0352	0.003
	Maths	5.863	0.00015

Table 13: One-sided ANOVAs of the imprecision related to an event (dependent variable) on the event (factor), for each source. (H_0 : the imprecision is identical across all events in the source.)

1155 B.2. Bayesian estimation

1156 B.2.1. Statistical approach

1157 **Estimation of upper and lower CDFs in EXP 1 and EXP 2** Recall that T denotes the
1158 space of possible values of the variables of interest (minimum temperatures in EXP 1, grades in
1159 EXP 2). For each source, we estimate general models of the form:

$$\begin{cases} \underline{p}(E) = \underline{f}(E) + \underline{\epsilon} \\ \bar{p}(E) = \bar{f}(E) + \bar{\epsilon} \end{cases} \quad (17)$$

1160 where $\underline{p}(E)$ (resp. $\bar{p}(E)$) are the elicited lower (resp. upper) probabilities of the cumulative
1161 event E (Section 2.2), \underline{f} and \bar{f} are CDFs over T from specified two-parameter families, with
1162 parameters $\underline{a}, \underline{b}$ (resp. \bar{a}, \bar{b}), and $\underline{\epsilon}$ and $\bar{\epsilon}$ are zero-mean normal distributions with variance $\underline{\sigma}^2$
1163 and $\bar{\sigma}^2$ respectively.

1164 For each equation, the parameter space is $\Theta \subseteq \mathbb{R}^3$, with a typical point $(\underline{a}, \underline{b}, \underline{\sigma})$ (resp.
1165 $(\bar{a}, \bar{b}, \bar{\sigma})$) specifying an \underline{f} (resp. \bar{f}) and the variance of the relevant error term. We specify
1166 the following priors over the hyperparameters : a, b, σ are realisations from $A \sim N(\mu_a, \sigma_a^2)$,
1167 $B \sim N(\mu_b, \sigma_b^2)$ and $\Sigma = \sigma_\sigma | Y |$ with $Y \sim N(0, 1)$.

1168 We use a MCMC-like approach to estimate the posterior distributions of these distributions
1169 through the use of the Python package PyMC3, and more specifically, the No-U-Turn Sampler
1170 algorithm (NUTS) (Hoffman and Gelman, 2014).

1171 The likelihood of observations x_1, \dots, x_n pertaining to t_1, \dots, t_n (e.g. elicited lower probab-
1172 ities for cumulative events $E_{t_i} = \{t \in T : t \leq t_i\}$) given the point $(a, b, \sigma) \in \Theta$ is:

$$L(a, b, \sigma | x_1, \dots, x_n) = \prod_{i \in \{1, \dots, n\}} \varphi \left(\frac{x_i - f_{(a,b)}(\{t \leq t_i\})}{\sigma} \right)$$

1173 where $f_{(a,b)}$ is the CDF with parameters a, b and φ is the density of the normal distribution.

	Temperature (EXP 1)	Grade (EXP 2)
Family 1	Truncated Normal $\mathcal{N}(a, b)$	Truncated Normal $\mathcal{N}(a, b)$
Family 2	Beta $B(a, b)$	Beta $B(a, b)$
Support	[min of min stated temperature, max of max stated temperature]	[0,20]

Table 14: Families of distributions over T (temperature; mark)

Note the minima and maxima in the first column are taken across all subjects' responses (Section 3.3).

Hence the likelihood of hyperparameters $\mu_a, \sigma_a^2, \mu_b, \sigma_b^2, \mu_\sigma, \sigma_\sigma^2$ given observations $x_1 \dots x_n$ is :

$$L(\mu_a, \sigma_a^2, \mu_b, \sigma_b^2, \mu_\sigma, \sigma_\sigma^2 | x_1, \dots, x_n) = \int_{(a,b,\sigma) \in \Theta} L(a, b, \sigma | x_1, \dots, x_n) dp(a, b, \sigma | \mu_a, \sigma_a^2, \mu_b, \sigma_b^2, \mu_\sigma, \sigma_\sigma^2)$$

$L(\mu_{\underline{a}}, \sigma_{\underline{a}}^2, \mu_{\underline{b}}, \sigma_{\underline{b}}^2, \mu_{\underline{\sigma}}, \sigma_{\underline{\sigma}}^2 | \underline{x}_1, \dots, \underline{x}_n)$ and $L(\mu_{\overline{a}}, \sigma_{\overline{a}}^2, \mu_{\overline{b}}, \sigma_{\overline{b}}^2, \mu_{\overline{\sigma}}, \sigma_{\overline{\sigma}}^2 | \overline{x}_1, \dots, \overline{x}_n)$ are used by the NUTS algorithm to estimate the posterior distributions of A, B and Σ , where $\underline{x}_1, \dots, \underline{x}_n, \overline{x}_1, \dots, \overline{x}_n$ are the elicited lower and upper probabilities respectively, under the parametric families for f given in Table 14.

Likelihood estimation of α in EXP 1 (Paris treatment) For the Bayesian estimation of the mixture coefficient α in the α -maxmin EU model, we supplement the general model (17) with the following equations

$$\begin{cases} MP(E) = \alpha \underline{p}(E) + (1 - \alpha) \overline{p}(E) + \epsilon_\alpha \\ 1 - MP(E^c) = \alpha \overline{p}(E) + (1 - \alpha) \underline{p}(E) + \epsilon_{\overline{\alpha}} \end{cases} \quad (18)$$

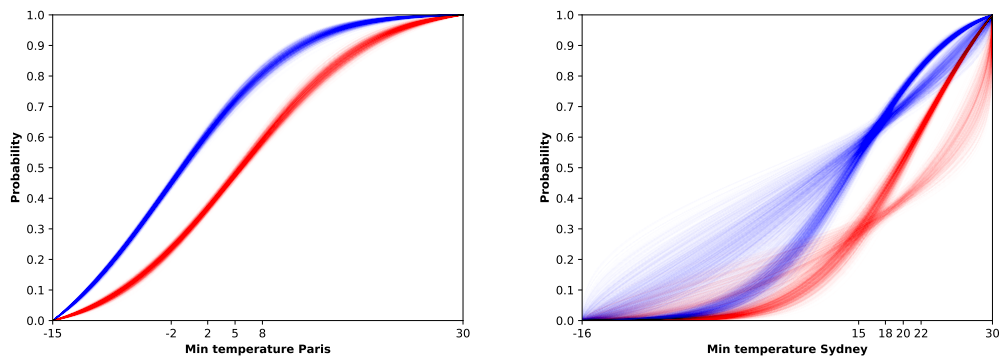
which are discussed in Section 4.5. We assume that α follows a beta distribution $B(a_\alpha, b_\alpha)$, and the $\epsilon_{\overline{\alpha}}$ and ϵ_α are zero-mean normal distributions, with the hyperparameters independent and normally distributed, as above, with variances $\sigma_{\overline{\alpha}}^2$ and σ_α^2 .

The MPs have been elicited for the Paris treatment in EXP 1. The hyperparameters concerning the upper and lower CDFs discussed above and those for α were estimated under the model composed of (17) and (18) using the NUTS algorithm, with the procedure set out above.

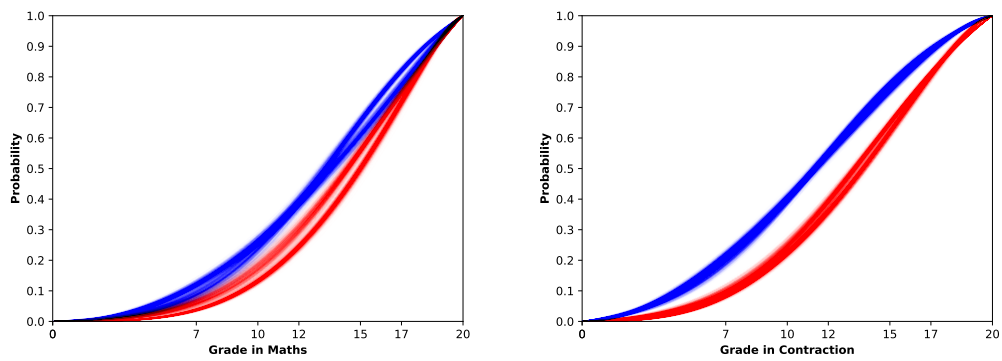
B.2.2. Analysis

Figure 15 displays the upper and lower distributions under the parametric families not shown in Figure 4. Tables 15-22 give statistics on the distribution over parameters under the estimated hyperparameters.

1192



(a) EXP 1



(b) EXP 2

Figure 15: Bayesian estimation of lower and upper CDFs: plots of 1000 samples from MCMC (Truncated Normal distribution for EXP 1; Beta distribution for EXP 2)

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
\underline{a}	5.25	0.31	0.01	4.65	5.87	1009.11	1.0
\bar{a}	-2.57	0.37	0.01	-3.27	-1.87	542.99	1.0
a_α	3.42	1.62	0.06	0.53	6.38	630.90	1.0
b_α	1.80	1.04	0.05	0.10	3.80	429.61	1.0
\underline{b}	11.35	0.75	0.02	9.71	12.66	1214.33	1.0
\bar{b}	11.00	0.64	0.03	9.78	12.17	614.95	1.0
$\underline{\sigma}$	0.22	0.01	0.00	0.20	0.23	1043.67	1.0
$\bar{\sigma}$	0.18	0.01	0.00	0.17	0.19	1055.83	1.0
$\sigma_{\bar{\alpha}}$	0.21	0.01	0.00	0.20	0.23	930.43	1.0
$\sigma_{\underline{\alpha}}$	0.19	0.01	0.00	0.17	0.20	909.78	1.0
α	0.81	0.04	0.00	0.74	0.88	754.54	1.0

Table 15: Statistics for parameters under Bayesian estimation; Paris (EXP 1); Normal parametrisation

Note mc_error: Monte Carlo procedure standard error; hdp_2.5 / hdp_97.5: Highest posterior density 2.5 and 97.5 percentiles; n_eff : count of iteration in the MCMC procedure.

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
\underline{a}	1.43	0.23	0.01	1.04	1.91	569.30	1.0
\underline{b}	1.64	0.29	0.01	1.13	2.24	571.99	1.0
\bar{a}	1.07	0.16	0.01	0.73	1.39	541.73	1.0
\bar{b}	2.46	0.35	0.01	1.76	3.17	522.49	1.0
a_α	4.32	1.76	0.06	1.06	7.87	812.98	1.0
b_α	1.90	1.08	0.04	0.18	3.92	606.36	1.0
$\underline{\sigma}$	0.22	0.01	0.00	0.20	0.23	1163.74	1.0
$\bar{\sigma}$	0.18	0.01	0.00	0.17	0.19	1239.11	1.0
$\sigma_{\bar{\alpha}}$	0.21	0.01	0.00	0.20	0.23	1134.15	1.0
$\sigma_{\underline{\alpha}}$	0.19	0.01	0.00	0.17	0.20	1409.46	1.0
α	0.81	0.04	0.00	0.74	0.88	1079.99	1.0

Table 16: Statistics for parameters under Bayesian estimation; Paris (EXP 1); Beta parametrisation

Note mc_error: Monte Carlo procedure standard error; hpd_2.5 / hpd_97.5: Highest posterior density 2.5 and 97.5 percentiles; n_eff : count of iteration in the MCMC procedure.

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
\underline{a}	22.03	0.45	0.01	21.11	22.80	1130.92	1.0
\bar{a}	14.66	0.46	0.01	13.78	15.48	876.71	1.0
\underline{b}	9.62	0.95	0.03	7.88	11.67	1018.57	1.0
\bar{b}	9.04	0.85	0.02	7.42	10.78	882.98	1.0
$\underline{\sigma}$	0.26	0.01	0.00	0.23	0.28	933.24	1.0
$\bar{\sigma}$	0.25	0.01	0.00	0.23	0.27	831.75	1.0

Table 17: Statistics for parameters under Bayesian estimation; Sydney (EXP 1); Normal parametrisation

Note mc_error: Monte Carlo procedure standard error; hpd_2.5 / hpd_97.5: Highest posterior density 2.5 and 97.5 percentiles; n_eff : count of iteration in the MCMC procedure.

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
\underline{a}	1.12	0.55	0.04	0.31	2.21	1.91	1.57
\bar{a}	0.14	0.32	0.03	-0.27	0.67	1.07	4.19
\underline{b}	1.32	0.48	0.02	0.49	2.24	320.38	1.00
\bar{b}	0.94	0.24	0.01	0.50	1.37	321.66	1.00
$\underline{\sigma}$	0.25	0.01	0.00	0.23	0.27	522.03	1.00
$\bar{\sigma}$	0.24	0.01	0.00	0.22	0.26	671.56	1.00

Table 18: Statistics for parameters under Bayesian estimation; Sydney (EXP 1); Beta parametrisation

Note mc_error: Monte Carlo procedure standard error; hpd_2.5 / hpd_97.5: Highest posterior density 2.5 and 97.5 percentiles; n_eff : count of iteration in the MCMC procedure.

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
\underline{a}	15.88	0.17	0.01	15.57	16.22	788.64	1.0
\bar{a}	5.40	0.28	0.01	4.86	5.95	955.04	1.0
\underline{b}	13.97	0.17	0.00	13.65	14.33	1218.24	1.0
\bar{b}	5.03	0.25	0.01	4.58	5.53	928.02	1.0
$\underline{\sigma}$	0.14	0.01	0.00	0.13	0.16	1158.52	1.0
$\bar{\sigma}$	0.17	0.01	0.00	0.16	0.19	958.54	1.0

Table 19: Statistics for parameters under Bayesian estimation; Maths (EXP 2); Normal parametrisation

Note mc_error: Monte Carlo procedure standard error; hpd_2.5 / hpd_97.5: Highest posterior density 2.5 and 97.5 percentiles; n_eff : count of iteration in the MCMC procedure.

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
\underline{a}	3.76	0.09	0.0	3.58	3.93	978.44	1.0
\bar{a}	1.46	0.05	0.0	1.36	1.55	900.67	1.0
\underline{b}	2.27	0.09	0.0	2.10	2.43	1049.31	1.0
\bar{b}	1.22	0.05	0.0	1.12	1.32	1031.36	1.0
$\underline{\sigma}$	0.16	0.01	0.0	0.14	0.17	1311.22	1.0
$\bar{\sigma}$	0.19	0.01	0.0	0.17	0.20	1247.27	1.0

Table 20: Statistics for parameters under Bayesian estimation; Maths (EXP 2); Beta parametrisation

Note mc_error: Monte Carlo procedure standard error; hpd_2.5 / hpd_97.5: Highest posterior density 2.5 and 97.5 percentiles; n_eff : count of iteration in the MCMC procedure.

1193

1194

1195

1196 **B.3. Matching Probability data and analysis of α**

1197 Table 23 provides descriptive statistics on the elicited MPs. Table 24 provide descriptive statis-
1198 tics on the α estimated from the raw data (from equations (9) and (10)). These equations cannot
1199 be applied to estimate α whenever the upper and lower probabilities of an event coincide, i.e.
1200 $\underline{p}(E) = \bar{p}(E)$; Table 24 performs the estimates using all events for which the equations can be
1201 applied—and hence only removes the two subjects for which the upper and lower probabilities
1202 coincide for all events (Table 12).

1203 **B.4. Elicitation-free check of $\alpha > \frac{1}{2}$**

1204 Under the α -maxmin EU model (1), it follows from Eqs. 9 and 10 that

$$MP(E) + MP(E^c) = 1 + (\bar{p}(E) - \underline{p}(E)) \cdot (1 - 2\alpha)$$

1205 Since $\bar{p}(E) - \underline{p}(E) \geq 0$ by definition, it follows that, whenever there is imprecision, $MP(E) +$
1206 $MP(E^c) < 1$ if and only if $\alpha > \frac{1}{2}$.

1207 Table 25 displays the descriptive statistics for the sum $MP(E) + MP(E^c)$ for the Paris
1208 source in EXP1. It is clear that the vast majority of subjects have a sum of MPs less than 1
1209 indicating an α greater than 0.5. Indeed, over 80% of subjects have sum of MPs less than or
1210 equal to 1.

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
\underline{a}	14.17	0.13	0.00	13.93	14.42	1090.28	1.0
\bar{a}	5.32	0.22	0.01	4.90	5.78	1068.01	1.0
\underline{b}	11.61	0.13	0.00	11.37	11.87	1167.95	1.0
\bar{b}	5.39	0.21	0.01	4.99	5.78	1209.38	1.0
$\underline{\sigma}$	0.12	0.01	0.00	0.11	0.13	1501.55	1.0
$\bar{\sigma}$	0.12	0.01	0.00	0.11	0.13	1144.85	1.0

Table 21: Statistics for parameters under Bayesian estimation; Contraction (EXP 2); Normal parametrisation

Note mc_error: Monte Carlo procedure standard error; hpd_2.5 / hpd_97.5: Highest posterior density 2.5 and 97.5 percentiles; n_eff : count of iteration in the MCMC procedure.

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
\underline{a}	3.06	0.09	0.0	2.90	3.23	888.13	1.00
\bar{a}	1.57	0.05	0.0	1.48	1.67	937.24	1.00
\underline{b}	1.96	0.07	0.0	1.82	2.11	478.84	1.01
\bar{b}	1.56	0.06	0.0	1.45	1.66	460.19	1.00
$\underline{\sigma}$	0.14	0.01	0.0	0.13	0.15	966.76	1.00
$\bar{\sigma}$	0.13	0.01	0.0	0.12	0.15	894.44	1.00

Table 22: Statistics for parameters under Bayesian estimation; Contraction (EXP 2); Beta parametrisation

Note mc_error: Monte Carlo procedure standard error; hpd_2.5 / hpd_97.5: Highest posterior density 2.5 and 97.5 percentiles; n_eff : count of iteration in the MCMC procedure.

$MP(E_t)$	count	mean	std	min	25%	50%	75%	max
Event $t =$								
-2	80	0.35	0.21	0.02	0.17	0.37	0.47	1.00
2	80	0.44	0.20	0.02	0.27	0.47	0.57	0.97
5	80	0.54	0.23	0.02	0.37	0.55	0.68	0.97
8	80	0.60	0.21	0.17	0.47	0.57	0.76	0.97

$1 - MP(E_t^c)$	count	mean	std	min	25%	50%	75%	max
Event $t =$								
-2	80	0.50	0.19	0.03	0.38	0.48	0.63	0.98
2	80	0.59	0.19	0.23	0.48	0.57	0.74	0.98
5	80	0.71	0.20	0.23	0.53	0.73	0.92	0.98
8	80	0.77	0.17	0.43	0.63	0.80	0.93	0.98

Table 23: Descriptive statistics for $MP(E_i)$ and $1 - MP(E_i^c)$ in Paris treatment, EXP 1

α	
count	78
mean	0.97
std	0.66
min	-0.32
25%	0.62
50%	0.80
75%	1.17
max	3.84

Table 24: Descriptive statistics for α , estimated from raw data according to Eqns (9) and (10). Estimation conducted across all subjects such that, for any least one event E , $\underline{p}(E) \neq \bar{p}(E)$.

$MP(E_t) +$ $MP(E_t^c)$	count	mean	std	min	25%	50%	75%	max
Event $t =$								
-2	80	0.84	0.20	0.29	0.71	0.89	0.98	1.31
2	80	0.85	0.20	0.29	0.73	0.89	0.99	1.34
5	80	0.83	0.22	0.24	0.69	0.89	0.99	1.29
8	80	0.83	0.18	0.39	0.69	0.89	0.99	1.26

Table 25: Empirical distribution of average $MP(E) + MP(E^c)$ across all events for which MPs were elicited (those concerning Paris temperature in EXP1).

1211 C. Experimental design and displays

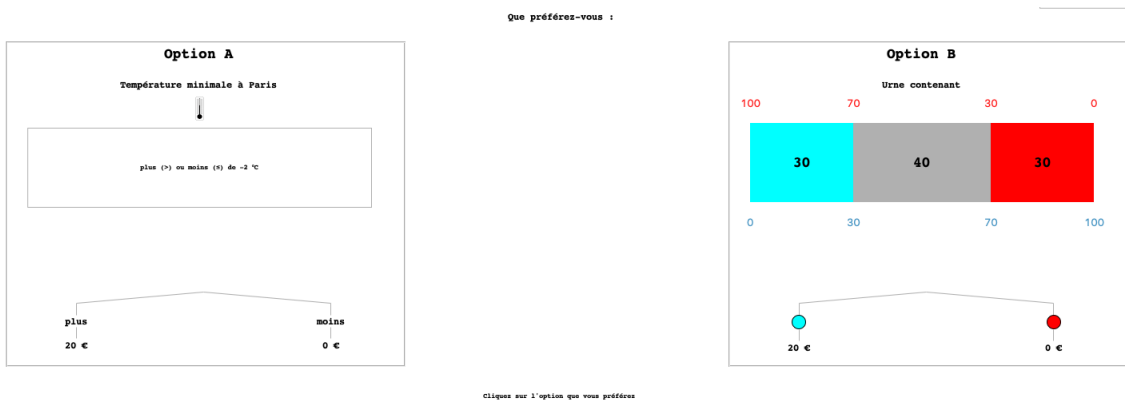
1212 C.1. Probability interval elicitation: displays

1213 Figure 16 shows the display in a typical step of the binary-choice procedure. Specifically, the
1214 two figures show the two choice questions making up the step, involving bets on complementary
1215 events (temperature below vs above; bet on red vs blue).

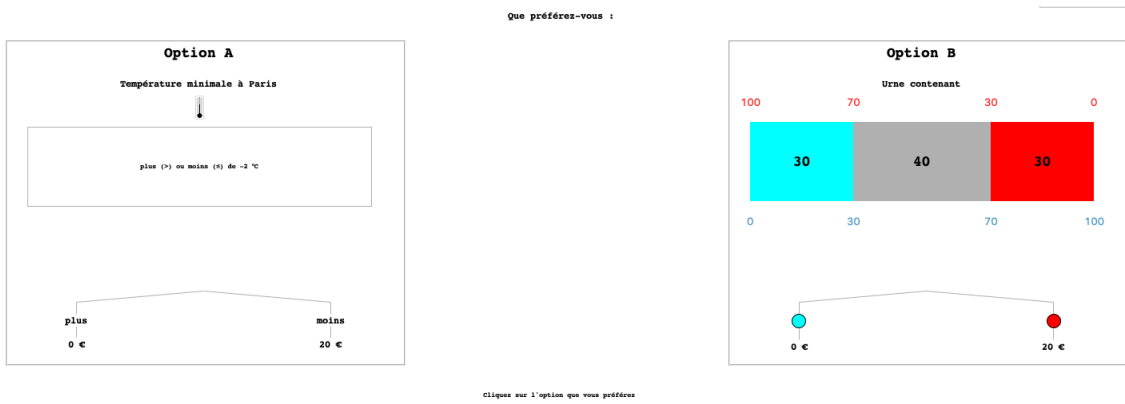
1216 At the end of the binary choice procedure, the two-cursor scrollbar, realising the 2D-choice
1217 list described in Section 2.5, is displayed, and the subject is invited to verify all choices, and
1218 correct them if required, prior to confirmation. The top pane of Figure 17 shows a typical
1219 confirmation screen that appears at the end of the binary-choice procedure, where the retained
1220 values for red and blue balls are 66 and 29 respectively. The red lines below then above the bar
1221 indicate that, for an urn with at least 29 blue balls and a minimum number of red balls greater
1222 than 66, option B (the bet on red from the urn) is preferred over A (the bet on the temperature
1223 being less than -2°C), whereas when there are at least 29 blue urns and the minimum number
1224 of red balls is less than 66, option A is preferred over B. Similarly, the blue lines indicate how
1225 preferences vary over urns with at least 66 red balls, for different minimum numbers of blue
1226 balls.

1227 The bottom pane of Figure 17 illustrates a situation where the subject is verifying and
1228 correcting preferences as the number of blue balls vary, by moving the blue cursor (which is
1229 thus highlighted). The red cursor is kept fixed at its provisional value,²¹ and, for each position
1230 of the blue cursor, the choice between the bet on the event (temperature greater than 2°) and
1231 the bet on the urn with the specified minimum number of blue balls and at least 66 red balls
1232 is presented, with the chosen option (as per Section 2.5) indicated. In this example, whilst the
1233 provisional values imply a preference for the bet on the event over that on the urn when only

²¹If the subject tries to move the red cursor, the blue cursor returns to its provisional value, and remains fixed there whilst the red cursor is being moved.

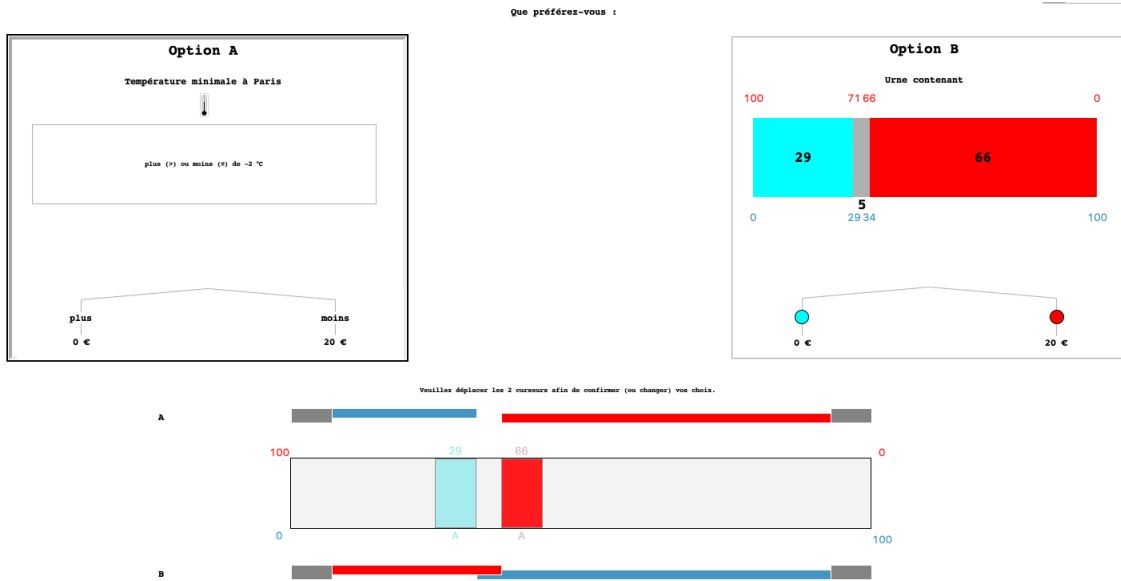


(a) Bet on the complement of the event E_{-2}

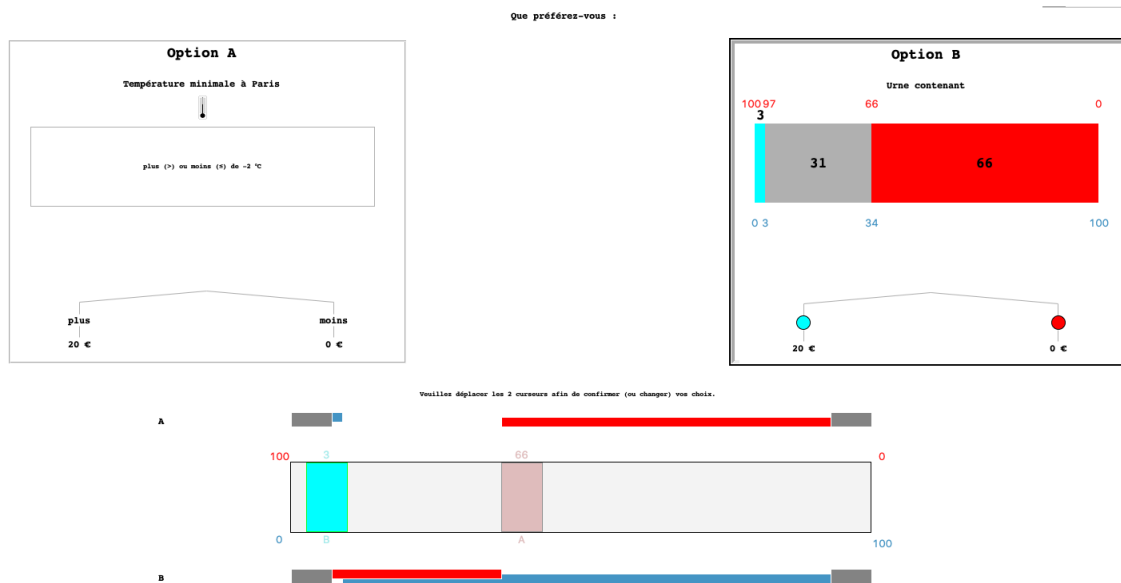


(b) Bet on the event E_{-2}

Figure 16: Displays for a step in the binary-choice procedure



(a) First display



(b) Moving the blue cursor to confirm and correct choices

Figure 17: Two-cursor scrollbar confirmation screen, implementing the 2D choice list

1234 3 balls are guaranteed to be blue, the subject prefers the urn. She may change her choice by
 1235 clicking on the urn or on the cursor. The bottom pane of Figure 17 shows the display just after
 1236 she has corrected her choice: the highlighted preferred option is now the bet on the urn, and the
 1237 blue lines above and below the scrollbar are modified accordingly. To confirm her response for
 1238 the event, the subject has to scroll the blue cursor across the entire confirmation line, scanning
 1239 all the choices, and likewise for the red cursor.

1240 In EXP 2, there was a final confirmation screen after the elicitation for all events in a given
 1241 source, presented in Figure 18. All interval-valued urns corresponding to the choices made and
 1242 confirmed by the subject for the source are presented on the left. They are graphically depicted
 1243 on the right: the red line shows the minimum number of red balls for each event (mark, in the
 1244 case of this source), whereas the blue line plots 100 minus the minimum number of blue balls.
 1245 To change a choice, a subject can either click on the choice on the right hand plot or on the
 1246 corresponding urn in the sidebar on the left. By doing so, she returns to the corresponding
 1247 two-cursor scrollbar confirmation screen, as in Figure 17. She may modify her choices on this
 1248 screen as described above, and must reconfirm before proceeding.

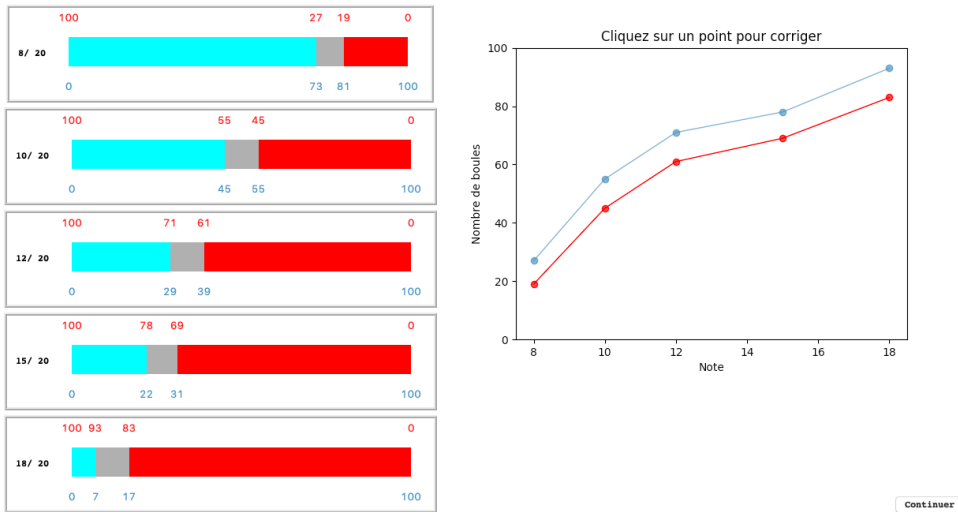
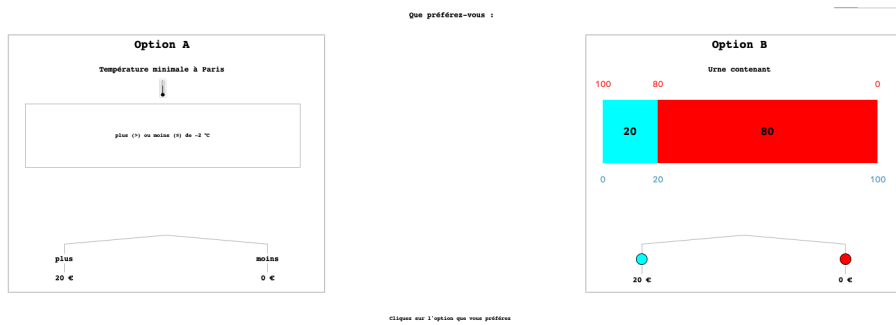


Figure 18: Omnibus confirmation screen

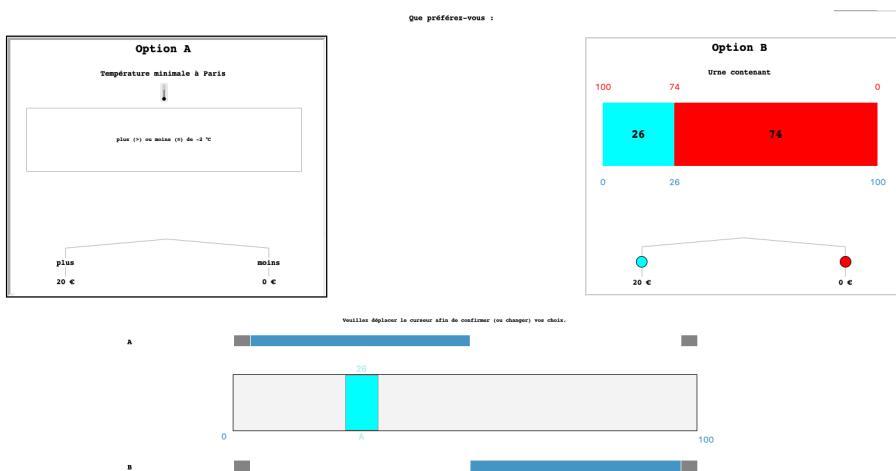
1249

1250 C.2. MP elicitation: displays

1251 Figure 19 shows the displays for a typical choice in the MP elicitation (top pane) and the
 1252 confirmation screen (bottom pane). These are comparable to the displays for probability interval
 1253 elicitation, with the exception that a standard single-cursor scrollbar is used for confirmation.
 1254 For the latter, as for the probability interval confirmation screen, the subject may use the cursor
 1255 to scan choices and may click on the relevant option to modify her choice. She must scan all
 1256 choices before confirming.



(a) Binary-choice procedure



(b) Confirmation scrollbar

Figure 19: MP displays

1257 **D. Technical Appendix: Lemmas for the proof of Proposition 4**

1258 In the following Lemmas, we suppose that preferences are represented according to (1) with
 1259 $\alpha > \frac{1}{2}$, with E the event of interest with the subjective probability interval $[\underline{p}(E), \bar{p}(E)]$.

1260 **Lemma 1.** *For every $m \in [0, 0.5]$:*

- 1261 • *If $\sigma_1([\underline{p}(E), \bar{p}(E)]) < m$, there exists $[p, q] \in R - B$ with $\sigma_1([p, q]) = m$, but no $[p, q] \in W$*
 1262 *with $\sigma_1([p, q]) = m$;*
- 1263 • *If $\sigma_1([\underline{p}(E), \bar{p}(E)]) > m$, there exists $[p, q] \in W$ with $\sigma_1([p, q]) = m$, but no $[p, q] \in R - B$*
 1264 *with $\sigma_1([p, q]) = m$;*
- 1265 • *If $\sigma_1([\underline{p}(E), \bar{p}(E)]) = m$, each $[p, q]$ with $\sigma_1([p, q]) = m$ and $[p, q] \neq [\underline{p}(E), \bar{p}(E)]$ belongs*
 1266 *to either R or B .*

1267 *Proof.* Straightforward to check from the representation (1) and the definition of σ (12). (See
 1268 also Figures 1 and 9.) □

1269 **Lemma 2.** *If $\underline{p}(E) = \bar{p}(E)$ and Case 1 arrives at a point $[p_i, q_i]$ with $p_i = q_i$, then the procedure*
 1270 *remains in Case 1, and $[\underline{p}_n, \bar{p}_n] \rightarrow [\underline{p}(E), \bar{p}(E)]$ as $n \rightarrow \infty$.*

1271 *Proof.* Once the procedure reaches the subspace of precise probabilities, it executes a standard
 1272 bisection procedure (lines 13–16, Figure 11). □

1273 **Lemma 3.** *Suppose that the procedure reaches a point $[p_i, q_i]$ in $R-B$ or W . Then the sequence*
 1274 *$\sigma_1([p_n, q_n]) \rightarrow \sigma_1([\underline{p}(E), \bar{p}(E)])$ as $n \rightarrow \infty$.*

Proof. Consider a stage i in the procedure where a point has just been found in $R-B$ or W . So
 the area containing the MPI is Φ_i (Eq. (14)). Let

$$\begin{aligned}
 m_i^W &= \min \sigma_1(\Phi_i) \\
 &= \sigma_1([\max \{p' : [p', q'] \in El_n \cap W\}, \min \{q' : [p', q'] \in El_n \cap W\}]) \\
 m_i^{RB} &= \max \sigma_1(\Phi_i) \\
 &= \begin{cases} \sigma_1([\min \{p'' : [p'', q''] \in El_n \cap R - B\}, \max \{q'' : [p'', q''] \in El_n \cap R - B\}]) & El_n \cap R - B \neq \emptyset \\ 0.5 & \text{otherwise} \end{cases}
 \end{aligned}$$

1275 and $|\Phi_i| = m_i^{R-B} - m_i^W$. The latter is the maximum difference in σ_1 values across all pairs
 1276 of points in Φ_i . In the first two subcases of Case 3 (lines 35-39), the next probability interval
 1277 elicited is

$$\begin{aligned}
 [p_{i+1}, q_{i+1}] &= \frac{1}{2}[\max \{p' : [p', q'] \in El_n \cap W\}, \min \{q' : [p', q'] \in El_n \cap W\}] \\
 &\quad + \frac{1}{2}[\min \{p'' : [p'', q''] \in El_n \cap R - B\}, \max \{q'' : [p'', q''] \in El_n \cap R - B\}]
 \end{aligned}$$

1278 In the second subcase of Case 2 (lines 22-23), where a point in W has been found, but no point
 1279 in R-B, the next probability interval elicited is

$$[p_{i+1}, q_{i+1}] = \frac{1}{2} \left[\frac{1}{2} \left(\begin{array}{l} \min \{p'' : [p'', q''] \in El_n \cap W\} + \\ \max \{q'' : [p'', q''] \in El_n \cap W\} \end{array} \right), \frac{1}{2} \left(\begin{array}{l} \min \{p'' : [p'', q''] \in El_n \cap W\} + \\ \max \{q'' : [p'', q''] \in El_n \cap W\} \end{array} \right) \right] \\
 + \frac{1}{2} [\min \{p'' : [p'', q''] \in El_n \cap W\}, \max \{q'' : [p'', q''] \in El_n \cap W\}]$$

1280 where $[\frac{1}{2}(\min \{p'' : [p'', q''] \in El_n \cap R - B\} + \max \{q'' : [p'', q''] \in El_n \cap R - B\}), \frac{1}{2}(\min \{p'' : [p'', q''] \in El_n \cap R - B\} + \max \{q'' : [p'', q''] \in El_n \cap R - B\})$
 1281 is the point on the diagonal of precise probabilities (i.e. degenerate probability intervals) that is
 1282 closest to $[\min \{p'' : [p'', q''] \in El_n \cap W\}, \max \{q'' : [p'', q''] \in El_n \cap W\}]$ (it is on the downwards
 1283 sloping 45° line from $[\min \{p'' : [p'', q''] \in El_n \cap W\}, \max \{q'' : [p'', q''] \in El_n \cap W\}]$). So this
 1284 point has σ_1 -value 0.5.

1285 In first subcase of Case 2 (lines 20-21), where a point in R-B has been found, but
 1286 no point in W, the next probability interval elicited, $[p_{i+1}, q_{i+1}]$, is a $\frac{1}{2} - \frac{1}{2}$ mix of
 1287 $[\min \{p'' : [p'', q''] \in El_n \cap R - B\}, \max \{q'' : [p'', q''] \in El_n \cap R - B\}]$ with

$$\left\{ \begin{array}{l} \left[\begin{array}{l} \min \{p'' : [p'', q''] \in El_n \cap R - B\} + \\ \max \{q'' : [p'', q''] \in El_n \cap R - B\} - 1 \end{array} \right], 1 \\ 0, \left[\begin{array}{l} \min \{p'' : [p'', q''] \in El_n \cap R - B\} + \\ \max \{q'' : [p'', q''] \in El_n \cap R - B\} \end{array} \right] \end{array} \right. \begin{array}{l} \min \{p'' : [p'', q''] \in El_n \cap R - B\} \\ + \max \{q'' : [p'', q''] \in El_n \cap R - B\} \\ \min \{p'' : [p'', q''] \in El_n \cap R - B\} \\ + \max \{q'' : [p'', q''] \in El_n \cap R - B\} \end{array} \begin{array}{l} > 1 \\ \\ \leq 1 \end{array}$$

1288 which is the point on the upper boundary (with either lower bound for the proba-
 1289 bility interval 0 or upper bound 1) that is on the downwards sloping 45° line from
 1290 $[\min \{p'' : [p'', q''] \in El_n \cap R - B\}, \max \{q'' : [p'', q''] \in El_n \cap R - B\}]$. This point has σ_1 -value
 1291 0.

Clearly, in all cases, $m_i^W < \sigma_1([p_{i+1}, q_{i+1}]) < m_i^{RB}$. Moreover, by the rest of the subcases in
 Cases 2 & 3, if this point is not in R-B or W, all the subsequent points elicited will have the same
 σ_1 -value as $[p_{i+1}, q_{i+1}]$. And whenever a point in R-B is found, the next area containing the
 MPI, Φ_{i+1} , will have the same minimum σ_1 -value m_i^W , but its maximum value will be replaced
 by $\sigma_1([p_{i+1}, q_{i+1}])$. By Lemma 4, it follows that

$$|\Phi_i| \cdot \frac{m_i^W}{m_i^{RB} + m_i^W} \leq |\Phi_{i+1}| \\
 \leq |\Phi_i| \cdot \frac{1 - m_i^W}{(1 - m_i^{RB}) + (1 - m_i^W)}$$

1292 Similarly, whenever a point in W is found, the next area containing the MPI, Φ_{i+1} , will have
 1293 the same maximum σ_1 value m_i^{RB} , but its minimum value will be replaced by $\sigma_1([p_{i+1}, q_{i+1}])$,
 1294 whence

$$|\Phi_i| \cdot \frac{1 - m_i^{RB}}{(1 - m_i^{RB}) + (1 - m_i^W)} \leq |\Phi_{i+1}|$$

$$\leq |\Phi_i| \cdot \frac{m_i^{RB}}{m_i^{RB} + m_i^W}$$

1295 Since, for any $j > i$, $m_j^{RB} \leq m_i^{RB}$ and $m_j^W \geq m_i^W$, for any such j , $\frac{1 - m_j^W}{(1 - m_j^{RB}) + (1 - m_j^W)} \leq$
1296 $\frac{1 - m_i^W}{(1 - m_i^{RB}) + (1 - m_i^W)}$ and $\frac{m_j^{RB}}{m_j^{RB} + m_j^W} \leq \frac{m_i^{RB}}{m_i^{RB} + m_i^W}$. So, for any $j = i + k$ with $k \in \mathbb{N}$, $k \geq 1$,
1297 $|\Phi_j| \leq \left(\max \left\{ \frac{1 - m_i^W}{(1 - m_i^{RB}) + (1 - m_i^W)}, \frac{m_i^{RB}}{m_i^{RB} + m_i^W} \right\} \right)^k \cdot |\Phi_i|$. So the sequence $[m_n^W, m_n^{RB}]$ is a bisection-
1298 like sequence of decreasing intervals (in the sense of containment), each of which contains
1299 $\sigma_1([p(E), \bar{p}(E)])$. Moreover, by the previous observation, whenever a point $[p, q]$ is found
1300 in W with $\sigma_1([p, q]) > 0$, then the sequence $|\Phi_n| = m_n^{RB} - m_n^W \rightarrow 0$ as $n \rightarrow \infty$, so
1301 $\sigma_1([p_n, q_n]) \rightarrow \sigma_1([p(E), \bar{p}(E)])$ as required. (Recall that $0.5 \geq m_n^{RB} \geq m_n^W \geq 0$ for all n .)

1302 We now separate two cases, according to whether $\sigma_1([p(E), \bar{p}(E)]) = 0$ or not. Suppose
1303 first that $\sigma_1([p(E), \bar{p}(E)]) = \delta > 0$. We show that the procedure will either arrive at a point
1304 with σ_1 -value δ , or a point in W . At a stage i in the procedure where no points in W have
1305 been found, but a point in R-B has, $m_i^W = 0$ and $0.5 \geq m_i^{R-B} > 0$. At each subsequent
1306 stage, by Lemma 1, either i. no point is found in W or R-B; ii. a point is found in W or
1307 R-B, and the next such point is in W ; iii. a point is found in W or R-B, and the next such
1308 point is in R-B. In case ii., the claim is established; in case i., by Lemma 1, the procedure is
1309 examining points with σ_1 -value δ , and the claim is established. Assume for reductio that at all
1310 such stages, the σ_1 -value of the explored points is not δ , but no point in W is found—i.e. we
1311 are always in case iii. Then, by the previous observations, for every $j = i + k$ with $k \in \mathbb{N}$, $k \geq 1$,
1312 $|\Phi_j| \leq \left(\frac{1 - m_i^W}{(1 - m_i^{RB}) + (1 - m_i^W)} \right)^k \cdot |\Phi_i| = \left(\frac{1}{2 - m_i^{RB}} \right)^k \cdot m_i^{RB}$. Hence $|\Phi_j| = m_j^{RB} \rightarrow 0$, contradicting the
1313 fact that there are no points with σ_1 -value less than δ in R-B. Hence the procedure eventually
1314 finds a point in W . By the previous observation it follows that $\sigma_1([p_n, q_n]) \rightarrow \sigma_1([p(E), \bar{p}(E)])$
1315 as required.

1316 Now consider the case where $\sigma_1([p(E), \bar{p}(E)]) = 0$. By Lemma 1, whenever the procedure
1317 searches for a point on a line $\sigma_1([p, q]) = m > 0$, it will find a point in R-B. Hence, by the
1318 previous argument, it produces a sequence of points $[p_n, q_n]$ in R-B, defining Φ_n and associated
1319 $[m_n^W, m_n^{RB}]$, with $m_n^W = 0$ and $m_n^{RB} \rightarrow 0$, as required.

1320 □

1321 **Lemma 4.** *Let $[p_W, q_W]$ be a point in W , with $\sigma_1([p_W, q_W]) = m_W$ and suppose that the line*
1322 *$\sigma_1([p, q]) = m_{R-B}$ contains a point in R-B but not in W . Then, for any point $[p_{R-B}, q_{R-B}] \in$*
1323 *$R - B$ with $\sigma_1([p_{R-B}, q_{R-B}]) = m$*

$$\sigma_1\left(\left[\frac{p_W + p_{R-B}}{2}, \frac{q_W + q_{R-B}}{2}\right]\right) \in \left[\frac{2m_W \cdot m_{R-B}}{m_W + m_{R-B}}, \frac{m_W(1 - m_{R-B}) + m_{R-B}(1 - m_w)}{(1 - m_{R-B}) + (1 - m_w)}\right]$$

1324 *Moreover, the same holds for a given point $[p_{R-B}, q_{R-B}] \in R - B$ and any point $[p_W, q_W] \in W$*
1325 *on the line $\sigma_1([p, q]) = m_W$.*

1326 *Proof.* We first restrict attention to points $[p, q]$ with $p < 1 - q$ (or, in the polar-style coordinate
1327 system, $\alpha < \frac{1}{2}$). For any points $[p_1, q_1]$ and $[p_2, q_2]$, written in polar-style coordinate system as
1328 (m_1, α_1) and (m_2, α_2) , by (12) and (13), the midpoint (in Cartesian coordinates), $\frac{1}{2}[p_1, q_1] +$
1329 $\frac{1}{2}[p_2, q_2]$ is $\left(\frac{\alpha_1 m_1 + \alpha_2 m_2}{\alpha_1 + \alpha_2}, \frac{\alpha_1 + \alpha_2}{2}\right)$ in the polar system. Written in the polar coordinate system, let
1330 $[p_W, q_W]$ be (m_W, α_W) ; the points on the line $\sigma_1([p, q]) = m_{R-B}$ are (m_{R-B}, α) , for varying α .
1331 Note that, by Proposition 2, $m_{R-B} > m_W$. It follows from representation 1 that $(z, [p', q'], 0) \prec$
1332 $(z, [p, q], 0)$ whenever $q' < q$ and $p' < p$, whence, since $[p_W, q_W] \in W$, we have that $(z, [p', q'], 0) \prec$
1333 $(z, E, 0)$ for all $q' < q_W$ and $p' < p_W$, so such points are not in R-B. So any point $[p, q]$ on
1334 $\sigma_1([p, q]) = m_{R-B}$ which is in R-B is such that $p \geq p_W$. By a similar argument (using the fact
1335 that $(0, [p', q'], z) \prec (0, E, z)$ for all $q' > q_W$ and $p' > p_W$), any point $[p, q]$ on $\sigma_1([p, q]) = m_{R-B}$
1336 which is in R-B is such that $q \geq q_W$. So any point $[p, q]$ on $\sigma_1([p, q]) = m_{R-B}$ which is in
1337 R-B has $\alpha > \frac{\alpha_W m_W}{m_{R-B}}$ (where, by 13, this is in the α of the point on $\sigma_1([p, q]) = m_{R-B}$ with
1338 $p = p_W = 2\alpha_W m_W$); similarly, any such point has $\alpha < \frac{\alpha_W(1-m_W)}{(1-m_{R-B})}$. Plugging these bounds into
1339 the expression for the midpoint yields the result. Similar calculations yield the same result for the
1340 cases of $p > 1 - q$ for some or all of the point considered. Finally, analogous arguments establish
1341 the conclusion for $[p_{R-B}, q_{R-B}] \in R-B$ fixed and $[p_W, q_W] \in W$ on the line $\sigma_1([p, q]) = m_W$. \square

1342 References

- 1343 Abdellaoui, M., Baillon, A., Placido, L., and Wakker, P. P. (2011). The Rich Domain of Uncer-
1344 tainty: Source Functions and Their Experimental Implementation. *The American Economic*
1345 *Review*, 101(2):695–723.
- 1346 Abdellaoui, M., Bleichrodt, H., and L’Haridon, O. (2008). A tractable method to measure utility
1347 and loss aversion under prospect theory. *Journal of Risk and Uncertainty*, 36(3):245–266.
- 1348 Abdellaoui, M., Diecidue, E., Kemel, E., and Onculer, A. (2021). Temporal Risk Resolution:
1349 Utility versus Probability Weighting Approaches. *Management Science*.
- 1350 Abdellaoui, M., Vossman, F., and Weber, M. (2005). Choice-based Elicitation and Decom-
1351 position of Decision Weights for Gains and Losses under Uncertainty. *Management Science*,
1352 51:1384–1399.
- 1353 Anscombe, F. J. and Aumann, R. J. (1963). A Definition of Subjective Probability. *The Annals*
1354 *of Mathematical Statistics*, 34:199–205.
- 1355 Baillon, A. and Bleichrodt, H. (2015). Testing Ambiguity Models through the Measurement of
1356 Probabilities for Gains and Losses. *American Economic Journal: Microeconomics*, 7(2):77–
1357 100.
- 1358 Baillon, A., Bleichrodt, H., Keskin, U., l’Haridon, O., and Li, C. (2018a). The Effect of Learning
1359 on Ambiguity Attitudes. *Management Science*, 64(5):2181–2198.
- 1360 Baillon, A., Cabantous, L., and Wakker, P. P. (2012). Aggregating imprecise or conflicting
1361 beliefs: An experimental investigation using modern ambiguity theories. *Journal of Risk and*
1362 *Uncertainty*, 44(2):115–147.
- 1363 Baillon, A., Huang, Z., Selim, A., and Wakker, P. P. (2018b). Measuring ambiguity attitudes
1364 for all (natural) events. *Econometrica*, 86(5):1839–1858.
- 1365 Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*. Springer, New York.
- 1366 Berger, J. O., Insua, D. R., and Ruggeri, F. (2000). Bayesian robustness. *Robust Bayesian*
1367 *Analysis*, pages 1–32.
- 1368 Bewley, T. F. (2002). Knightian decision theory. Part I. *Decisions in Economics and Finance*,
1369 25(2):79–110.
- 1370 Borel, E. (1939). Valeur pratique et philosophie des probabilités.[Practical and philosophical
1371 value of probability]. *Traité du calcul des probabilités et de ses applications [Treatise of prob-*
1372 *ability calculus and its applications]*, 4.
- 1373 Bose, S. and Renou, L. (2014). Mechanism design with ambiguous communication devices.
1374 *Econometrica*, 82(5):1853–1872.

- 1375 Carney, M., Jon Cunliffe, Dave Ramsden, Andrew Haldane, Jonathan Haskel, Michael Saun-
1376 ders, Silvana Tenreyro, and Gertjan Vlieghe (2019). Inflation Report. *Bank of Eng-
1377 land*. [https://www.bankofengland.co.uk/-/media/boe/files/inflation-report/2019/
1378 may/inflation-report-may-2019.pdf](https://www.bankofengland.co.uk/-/media/boe/files/inflation-report/2019/may/inflation-report-may-2019.pdf).
- 1379 Chateauneuf, A., Eichberger, J., and Grant, S. (2007). Choice under uncertainty with the best
1380 and worst in mind: Neo-additive capacities. *Journal of Economic Theory*, 137(1):538–567.
- 1381 Chew and Sagi, J. S. (2006). Event exchangeability: Probabilistic sophistication without conti-
1382 nuity or monotonicity. *Econometrica*, 74(3):771–786.
- 1383 Chew, S. H., Miao, B., and Zhong, S. (2017). Partial ambiguity. *Econometrica*, 85(4):1239–1260.
- 1384 Clemen, R. T. and Reilly, T. (2013). *Making Hard Decisions with DecisionTools*. Cengage
1385 Learning.
- 1386 Cox, L. A. T. (2012). Confronting deep uncertainties in risk analysis. *Risk Analysis*, 32(10):1607–
1387 1629. <http://onlinelibrary.wiley.com/doi/10.1111/j.1539-6924.2012.01792.x/full>.
- 1388 Dieckmann, N. F., Mauro, R., and Slovic, P. (2010). The effects of presenting imprecise proba-
1389 bilities in intelligence forecasts. *Risk Analysis: An International Journal*, 30(6):987–1001.
- 1390 Dimmock, S. G., Kouwenberg, R., and Wakker, P. P. (2015). Ambiguity attitudes in a large
1391 representative sample. *Management Science*. [http://pubsonline.informs.org/doi/abs/
1392 10.1287/mnsc.2015.2198](http://pubsonline.informs.org/doi/abs/10.1287/mnsc.2015.2198).
- 1393 Eichberger, J., Grant, S., Kelsey, D., and Koshevoy, G. A. (2011). The α -MEU model: A
1394 comment. *Journal of Economic Theory*, 146(4):1684–1698.
- 1395 Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *Quart. J. Econ.*, 75:643–669.
- 1396 Epstein, L. G. and Schneider, M. (2010). Ambiguity and Asset Markets. *Annual Review of
1397 Financial Economics*, 2(1):315–346.
- 1398 Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., and Sunde, U. (2018). Global evidence
1399 on economic preferences. *The Quarterly Journal of Economics*, 133(4):1645–1692.
- 1400 Garlappi, L., Uppal, R., and Wang, T. (2007). Portfolio selection with parameter and model
1401 uncertainty: A multi-prior approach. *Review of Financial Studies*, 20(1):41–81. [http://rfs.
1402 oxfordjournals.org/content/20/1/41.short](http://rfs.oxfordjournals.org/content/20/1/41.short).
- 1403 Ghirardato, P., Maccheroni, F., and Marinacci, M. (2004). Differentiating ambiguity and ambi-
1404 guity attitude. *J. Econ. Theory*, 118(2):133–173.
- 1405 Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *J. Math.
1406 Econ.*, 18(2):141–153.

- 1407 Giustinelli, P., Manski, C. F., and Molinari, F. (2021). Precise or Imprecise Probabilities?
1408 Evidence from Survey Response on Late-onset Dementia. *Journal of the European Economic*
1409 *Association*.
- 1410 Grant, S., Rich, P., and Stecher, J. D. (2019). An Ordinal Theory of Worst- and Best-Case
1411 Expected Utility. SSRN Scholarly Paper ID 3369078, Social Science Research Network,
1412 Rochester, NY. <https://papers.ssrn.com/abstract=3369078>.
- 1413 Gul, F. and Pesendorfer, W. (2014). Expected uncertain utility theory. *Econometrica*, 82(1):1–
1414 39.
- 1415 Gul, F. and Pesendorfer, W. (2015). Hurwicz expected utility and subjective sources. *Journal*
1416 *of Economic Theory*, 159:465–488.
- 1417 Hey, J., Lotito, G., and Maffioletti, A. (2010). The descriptive and predictive adequacy of
1418 theories of decision making under uncertainty/ambiguity. *Journal of Risk and Uncertainty*,
1419 41(2):81–111. <http://dx.doi.org/10.1007/s11166-010-9102-0>.
- 1420 Hill, B. (2021). Beyond Uncertainty Aversion. HEC Paris Research Paper Series, HEC Paris.
- 1421 Hoffman, M. D. and Gelman, A. (2014). The No-U-Turn sampler: Adaptively setting path
1422 lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.*, 15(1):1593–1623.
- 1423 Ilut, C. L. and Schneider, M. (2014). Ambiguous business cycles. *American Economic Review*,
1424 104(8):2368–99.
- 1425 Johnson, C., Baillon, A., Bleichrodt, H., Li, Z., van Dolder, D., and Wakker, P. (2021). Prince:
1426 An Improved Method For Measuring Incentivized Preferences. *Journal of Risk and Uncer-*
1427 *tainty*.
- 1428 Karanki, D. R., Kushwaha, H. S., Verma, A. K., and Ajit, S. (2009). Uncertainty Analysis Based
1429 on Probability Bounds (P-Box) Approach in Probabilistic Safety Assessment. *Risk Analysis*,
1430 29(5):662–675.
- 1431 Karni, E. (2020). A mechanism for the elicitation of second-order belief and subjective informa-
1432 tion structure. *Economic Theory*, 69(1):217–232.
- 1433 Klibanoff, P., Mukerji, S., Seo, K., and Stanca, L. (2021). Foundations of ambiguity models
1434 under symmetry: α -MEU and smooth ambiguity. *Journal of Economic Theory*, page 105202.
- 1435 Kocher, M. G., Lahno, A. M., and Trautmann, S. T. (2018). Ambiguity aversion is not universal.
1436 *European Economic Review*, 101:268–283.
- 1437 Kriegler, E., Hall, J. W., Held, H., Dawson, R., and Schellnhuber, H. J. (2009). Imprecise
1438 probability assessment of tipping points in the climate system. *Proceedings of the national*
1439 *Academy of Sciences*, 106(13):5041–5046.

- 1440 Machina, M. J. and Schmeidler, D. (1992). A More Robust Definition of Subjective Probability.
1441 *Econometrica*, 60(4):745–780.
- 1442 Manski, C. F. (2003). *Partial Identification of Probability Distributions*. Springer Science &
1443 Business Media.
- 1444 Manski, C. F. (2013). *Public Policy in an Uncertain World : Analysis and Decisions*. Harvard
1445 University Press, Cambridge, Mass.
- 1446 Morgan, M. G. (2014). Use (and abuse) of expert elicitation in support of decision making for
1447 public policy. *Proceedings of the National Academy of Sciences*, 111(20):7176–7184. [http:](http://www.pnas.org/content/111/20/7176.short)
1448 [//www.pnas.org/content/111/20/7176.short](http://www.pnas.org/content/111/20/7176.short).
- 1449 Qiu, J. and Weitzel, U. (2016). Experimental evidence on valuation with multiple priors. *Journal*
1450 *of risk and uncertainty*, 53(1):55–74.
- 1451 Savage, L. J. (1954). *The Foundations of Statistics*. Dover, New York.
- 1452 Stoye, J. (2011). Axioms for minimax regret choice correspondences. *Journal of Eco-*
1453 *nomic Theory*, 146(6):2226–2251. [http://www.sciencedirect.com/science/article/pii/](http://www.sciencedirect.com/science/article/pii/S0022053111001372)
1454 [S0022053111001372](http://www.sciencedirect.com/science/article/pii/S0022053111001372).
- 1455 Wakker, P. P. (2010). *Prospect Theory For Risk and Ambiguity*. Cambridge University Press,
1456 Cambridge.
- 1457 Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London.