

# RANGES OF RANDOMIZATION\*

Marina Agranov<sup>†</sup>      Pietro Ortoleva<sup>‡</sup>

*This version:* September 23, 2020

Latest version available [here](#)

## Abstract

A growing literature has shown how people sometimes prefer to randomize between two options. We study how *prevalent* this behavior is in an experiment using a novel and simple method. We allow subjects to randomize between options in a series of questions in which one of the alternatives is fixed and the other varies, capturing the range of values for which subjects want to randomize. We find that most subjects choose to randomize in most questions. Crucially, they do so for ranges of values are ‘very large’: for example, when comparing a fixed amount  $x$  with a lottery that pays \$20 or \$0 with equal chances, subjects typically randomize *for all*  $x$ s between \$5.3 and \$12. Large ranges are found in other questions as well, showing how prevalent the desire to randomization is. We connect ranges to standard choices, Certainty-Bias, and non-Monotonicity.

**Key words:** Preference for Randomization, Incomplete Preferences, Non-Expected Utility.

**JEL:** C91, D81, D90

---

\*We thank Mark Dean, David Dillenberger, Efe Ok, and Leeat Yariv for useful comments and suggestions. Ortoleva gratefully acknowledges the financial support of NSF Grant SES-1763326.

<sup>†</sup>Division of the Humanities and Social Science, California Institute of Technology. Email: [magranov@hss.caltech.edu](mailto:magranov@hss.caltech.edu).

<sup>‡</sup>Department of Economics, Princeton University and Princeton School of Public and International Affairs. Email: [pietro.ortoleva@princeton.edu](mailto:pietro.ortoleva@princeton.edu).

# 1 Introduction

A large and growing literature documented how people sometimes prefer to *randomize* between options. For example, when asked to choose between a sure amount  $\$x$  and a lottery that pays  $\$20$  or  $\$0$  with equal chances, for some values of  $x$  the individual may prefer a randomization between the two alternatives over either of the two. Preferences for randomization have been documented in many environments and have been connected theoretically to violations of Expected Utility, convexity, and incomplete preferences. (Both the empirical and the theoretical literature are discussed below.)

To our knowledge, however, most studies focused on demonstrating the *existence* of choices in which subjects want to randomize. This paper studies their *prevalence*. For example, in the choice above between  $\$x$  and a lottery that pays  $\$20$  or  $\$0$ , we ask: How big is the *range* of values of  $x$  for which people wants to randomize? Is it large or is it small? We propose a novel, simple method to capture this range—a modified Multiple-Price-List—and we apply it to standard economic questions like certainty equivalents and lottery equivalents. We then relate our results to choices in standard tasks, Certainty Bias, non-Monotonocities, and other individual characteristics.

Understanding the prevalence of preferences for randomization is essential to assess their importance. If they exist solely on rare, smartly-constructed, knife-edge cases, they may be of theoretical interest but perhaps less consequential for many economic problems. If, instead, preferences for randomization apply more broadly, encompassing a variety of questions and values, then ignoring them can be problematic: if individuals do prefer to randomize, it should be taken into account in modeling and in interpreting data, e.g., for welfare calculations.

**Eliciting Ranges.** To illustrate our approach, recall a standard technique to elicit the certainty equivalent of a lottery, the Multiple Price List (MPL): a list of questions with the lottery fixed on the left, while on the right is an amount of dollars, increasing as we proceed down the rows. In each line, the agent has to choose either the left (lottery) or the right (money) option; the highest value against which the lottery is chosen indicates the certainty equivalent. We make a simple twist: in our experiment, instead of left or right, in each line subjects need to indicate a number between 0 and 10. If they choose 0, they get the left option; if 10, the right one; but 5 gives them each option with probability 50%, with a lottery run at the end; 3 means getting the left option with probability 30%, *etc.* That is, in each line subjects can choose left, right, or one of many lotteries between them.

Individuals who follows Expected Utility should never report numbers others than 0 or 10, except for one line if they are indifferent.<sup>1</sup> But others may prefer to randomize if they

---

<sup>1</sup>This is because, under Expected Utility,  $p \geq \delta_x \Leftrightarrow p \geq \lambda p + (1 - \lambda)\delta_x \geq \delta_x$  for all  $\lambda \in (0, 1)$ . In fact, this holds

violate Expected Utility and preferences have points of strict convexity. Documenting randomization thus implies documenting violations of Expected Utility and instances of strict convexity. As we discuss below, this can also be related to preferences incompleteness.

Our method allows us to elicit ranges of values in which subjects choose to randomize. We use it to capture ranges of i) certainty equivalents of a lottery, ii) lottery equivalents of a sure amount, and iii) lottery equivalents of another lottery. We also ask the same questions using standard MPLs, allowing a direct comparison; and we measure individual characteristics including risk attitudes, Certainty Bias, measures of IQ, and overconfidence.

**Results.** *Most subjects report ranges.* More than three-quarter of subjects report in at least two out of three questions ranges of values in which they want to randomize; only 16% never do.

Second, *ranges are 'very large'*. For example, when asked to choose between  $\$x$  for sure and  $\$20$  or  $\$0$  with equal chances, on average subjects want to randomize for all  $x$ s between  $\$5.3$  to  $\$12$ . That is, there is a large breadth of values, covering many possible risk attitudes, where subjects choose to randomize. Similar results hold for other questions. The same is true if we look at the number of rows where subjects choose to randomize: when subject randomize, on average they do so for about *half* of the rows in each question. Ranges remain common and sizable also if we take more restrictive definitions—for example, if we define a range as giving at least 40% of chances to both options.

Third, *ranges tend to involve values in the risk seeking domain*: for example, in the question above for 72% of subjects the top of the range exceeds the risk neutral value. We discuss how this is incompatible with risk aversion and signals more complex risk attitudes. In general, ranges encompass both risk averse and risk seeking values, although they are asymmetric and extend much further into risk averse areas.

Fourth, *choices in standard MPLs on average fall around the middle of ranges*. When taking more restrictive definition of ranges (at least 40% to both options), answers to standard MPLs fall instead towards more risk averse values.

Fifth, *ranges are related to Certainty-Bias and some measures of risk attitude*: less certainty-biased, and more risk averse subjects have more frequent and larger ranges.

Sixth, *virtually all subjects who exhibit non-Monotone choices in standard MPLs also have ranges*, and non-monotonicities and ranges are related in multiple ways. Instead of mistakes, non-monotonicities could be signs of more complex preferences.

Overall, our key finding is that the majority of our subjects exhibit preferences for randomization that span wide ranges encompassing values typically used in economic problems. Far from being restricted to knife-edge case, preferences for randomization

---

more generally than Expected Utility, only requiring Betweenness (Dekel, 1986).

are prevalent and widespread. Ignoring them in modeling or in interpreting data may substantially misrepresent choices and preferences.

**Related Literature.** A large theoretical literature discusses preferences from randomization (e.g., Machina 1985 and, more recently, Cerreia-Vioglio et al. 2019b); see Section 2 for a review. This is paired with a growing empirical literature that documents strict convexity of preferences under risk (Becker et al., 1964; Sopher and Narramore, 2000), randomization with objective lotteries (Agranov and Ortoleva, 2017; Dwenger et al., 2018), ambiguity (Cettolin and Riedl, 2019), time preferences (Agranov and Ortoleva, 2017), social preferences (Kircher et al., 2013; Agranov and Ortoleva, 2017; Miao and Zhong, 2018), and even choices with dominated options (Rubinstein, 2002). Agranov et al. (2020) documents high rates of preferences for randomization in different domains and show that most persist after explicit training, substantiating the belief that they represent a fundamental trait of preferences and not just an error. Closest to us are Permana (2020) and Feldman and Rehbeck (2020). The former allows subjects to choose “I am not sure what to choose” in MPLs, in which case subjects receive a random assignment.<sup>2</sup> Feldman and Rehbeck (2020) use the convex budgets method, in which subjects choose from budgets covering the space of three-outcome lotteries; they find that many subjects prefer non-degenerate mixtures.

Like these papers, we capture randomization using incentivized measures. But we give nuanced options of mixing and elicit the ranges of values for which randomization occurs.

Less directly, this paper is also related to the literature that aims to capture preference incompleteness or imprecision using variants of MPLs. The techniques used are reminiscent of ours but do not involve randomization and are not incentivized. Cohen et al. (1985, 1987) use MPLs in which in any row subjects are allowed to say they “do not know;” they find this is often used. But this is not incentivized, as payments are computed using as a switch point the middle of the interval in which the agent expressed a ranking. A separate group of papers measures preference imprecision—a notion distinct from, but related to, incomplete preferences: Dubourg et al. (1994, 1997); Butler and Loomes (2007, 2011); Cubitt et al. (2015). In these papers, subjects are either asked to choose between options, but also to report the strength of their preferences (which is inconsequential); or, they face a MPL in which they can choose not to make a choice in some rows, but also have to separately report their switching point (only the latter matters for payment.) These papers document sizable preferences imprecision and link it to features like violations of Independence and preference reversal (Butler and Loomes, 2007, 2011). More recently, Enke and Graeber (2019) measure “cognitive uncertainty:” in a series of questions including risk and ambiguity preferences, subjects first choose from a MPL; then, in a second screen

---

<sup>2</sup>The paper finds that this additional option is chosen not very frequently (less than 10% of all choices). This may be due to the non-neutral wording or to the fact that only the 50/50 mixture was offered.

they indicate two bounds they are “certain” about; these are inconsequential for payment, making these ranges not incentivized; they are also unrelated to randomization. They find sizable ranges and relate them to behavior in different areas, including risk and ambiguity. Like these papers, we also aim to measure ranges; but we do so using the desire to randomize (and thus have an incentive-based measure).

Finally, we relate ranges to violations of monotonicity in MPLs. We are not the first to suggest convexity as a possible explanation: Chew et al. (2019) document a link between violations of monotonicity in MPLs and deliberate randomization as well as other violations of Expected Utility. Our results are similar in spirit.

## 2 Theoretical Background

In our experiment subjects can report numbers between 0 and 10 to express their desire to randomize between options. What do theories predict?

**No Range with Expected Utility.** Agents whose choices maximize complete, monotone preferences following Expected Utility never choose numbers other than 0 or 10 in more than one row. Under Expected Utility, if  $p$  is strictly preferred to  $q$ , it is strictly preferred also to any mixture between them—an immediate consequence of the Independence axiom. It is only when  $p$  and  $q$  are indifferent that the agent is also indifferent with any randomization; since in MPLs one of the options is fixed and the other becomes strictly better, we can have indifference in at most one row. Thus, if we observe numbers other than 0 or 10 in more than one row, we are documenting a violation of complete preferences under Expected Utility.<sup>3</sup> This can be seen as related to non-Expected Utility and convexity of preferences; to incompleteness; or to other explanations. We discuss each below.

**non-Expected Utility and Convexity.** Suppose preferences are complete but allow them to violate Expected Utility. Choosing a number other than 0 or 10 means that the agent has a (weak) preference for convex combinations: for some  $\alpha \in (0, 1)$ ,  $\alpha p + (1 - \alpha)q \geq p, q$ . Many models allow for it. In Rank Dependent Expected Utility, or Cumulative Prospect Theory (Quiggin, 1982; Kahneman et al., 1991), preferences may be strictly convex if probability weighting is not always pessimistic. Under Cautious Expected Utility, preferences are in general convex (Cerreia-Vioglio et al., 2015); under the latter, however, there should not be instances of strict convexity with degenerate lotteries.<sup>4</sup> These also relate to interpretations of stochastic choice as the outcome of deliberate randomization, also due to strict convexity (Machina, 1985; Agranov and Ortoleva, 2017; Cerreia-Vioglio et al., 2019b).

<sup>3</sup>In fact, it is a violation of property weaker than Independence: Betweenness (Dekel, 1986).

<sup>4</sup>That is, we cannot have  $p, \delta_x, \alpha$  such that  $\alpha \delta_x + (1 - \alpha)p > p, \delta_x$ . This is directly implied by the Negative Certainty Independence axiom. See Cerreia-Vioglio et al. (2015, p. 697).

**Convexity and Risk Attitudes.** For any monetary lottery  $p$ , denote  $E[p]$  its expected value; denote degenerate lotteries that pay  $y$  by  $\delta_y$ . Recall that an agent is *risk averse*, if the expected value of a lottery is preferred to the lottery, i.e.,  $\delta_{E[p]} \succeq p$  for all  $p$ . In general, risk attitudes and convexity are unrelated: the former is linked to concavity of the utility over money (the Bernoulli index under Expected Utility), the latter is quasi-concavity in probabilities. It is easy to construct examples where the agent is risk averse or risk seeking yet weakly, and sometimes strictly, prefers to randomize.<sup>5</sup>

However, randomization can be informative about risk attitudes: if  $x > E[p]$ , randomizing between  $p$  and  $\delta_x$  is a *violation* of risk aversion. This is because choosing to randomize means  $q := \alpha p + (1 - \alpha)\delta_x \succeq \delta_x, p$  for some  $\alpha \in (0, 1)$ . If  $x > E[p]$ , then  $x > E[q]$ , thus  $\delta_x \succ \delta_{E[q]}$ . It follows that  $q \succ \delta_{E[q]}$ , violating risk aversion. This will be of particular relevance below.

**Incomplete Preferences.** An alternative approach is to consider evidence of randomization as evidence of *incompleteness*. A large and growing literature has studied incomplete preferences and their possible completions.<sup>6</sup> To see how randomization can be related to incompleteness, following Ghirardato et al. (2004); Cerreia-Vioglio (2010); Cerreia-Vioglio et al. (2015), for any preference relation  $\succeq$  over lotteries define its linear core  $\succeq'$  as its largest subrelation that satisfies independence, i.e.,  $p \succeq' q$  iff  $\lambda p + (1 - \lambda)r \succeq \lambda q + (1 - \lambda)r$  for all  $\lambda \in (0, 1)$  and lottery  $r$ . The relation  $\succeq'$  is naturally incomplete if  $\succeq$  violates Independence; the rankings it includes may be understood as those that the agent is sure about, as they are preserved when both options are mixed with others. The key observation here is that if a subject in our experiment prefers to randomize between  $p$  and  $q$ , then  $p$  and  $q$  must be *incomparable* according to the linear core  $\succeq'$ . Thus, evidence of randomization can be interpreted as evidence of incompleteness of preferences: the agent does not know what to choose and prefers to randomize. Our procedure therefore allows us to identify some points of incompleteness.<sup>7</sup>

**Preference Imprecision and Other approaches.** The ranges we elicit could be instead related to preference imprecision—distinct but connected to incomplete preferences. A formal treatment appear in Butler and Loomes (2011). Other recent theoretical models consider notions of cognitive noise or cognitive imprecision: see Khaw et al. (2017, 2018);

<sup>5</sup>For example, under Cautious Expected Utility, if all utilities in the representation are concave (resp. convex) then preferences are risk averse (seeking). Yet, these preferences are convex, and have points of strict convexity, for example, whenever there are finitely many utilities (Cerreia-Vioglio et al., 2019a).

<sup>6</sup>See, among many, Bewley (1986); Dubra et al. (2004); Ghirardato et al. (2004); Gilboa et al. (2010); Ok et al. (2012); Cerreia-Vioglio et al. (2015) and, more recently, Ok and Nishimura (2019) and references therein.

<sup>7</sup>Naturally, there are other forms of incompleteness that our procedure does not identify. It is also worth noting that incompleteness cannot be formally separated in our context from non-Expected Utility. For example, Cerreia-Vioglio et al. (2015) derived Cautious Expected Utility both starting from complete, non-EU preferences; and as completions of incomplete preferences.

Enke and Graeber (2019); Gabaix (2019) and references therein. While related to incompleteness, at least in their standard formulation none of these theories predict preference for randomization: including it would link these models to Cautious Expected Utility.

### 3 Design

The experiment included 10 main questions and several control tasks.<sup>8</sup> Main questions involved choices between monetary lotteries with objective probabilities. These questions were of two types: standard multiple price-lists (MPL, hereafter) and *range* multiple price-list (r-MPL, hereafter).

A standard MPL consists of several rows, each including two options, Left and Right. The Left option is the same in all rows, while the Right one changes, becoming more attractive as we go down the rows. Subjects are required to select one option in each row.

Range MPL, or r-MPL, are almost identical, but in each row subjects are asked, instead of Left or Right, to indicate an integer number from 0 to 10. This number corresponds to the probability of receiving the Left option: specifying an integer  $i$  means getting the Left option with probability  $10i\%$  and the Right option with the remaining probability of  $(100 - 10i)\%$ . Thus, indicating 3 means getting the Left option with probability 30%. This enriches the standard MPL technique by allowing subjects to choose any combination of two options while maintaining the ability to choose either option for sure.

Table 1 lists the 10 main questions, where the last column includes the range of values of the Right option (the exact steps are listed in Appendix A). Note that Q1r-2r-3r correspond to Q1-2-3, except that they use r-MPLs instead of MPLs. This allows us to compare behavior across types of questions. To investigate if responses are sensitive to specific kinds of questions, we purposely vary Left-Right combinations: fixed lottery vs. sure amount, fixed sure amount vs. lottery, and lottery vs. lottery. Questions Q4-Q7 measure risk attitudes and Certainty Bias—note that Q4 and Q5, and Q6 and Q7, have common-ratio-type variations.

In addition, subjects completed two investment tasks (Gneezy and Potters, 1997), measures of IQ (ICAR, Condon and Revelle, 2014) and overconfidence, and a non-incentivized questionnaire.<sup>9</sup> Appendix A.7 discusses the answer to this questionnaire finding them broadly consistent with choices.

Subjects' payment consisted of two parts in addition to the participation fee (\$7). First, one of the 10 main questions or two investment tasks was selected at random; if the chosen question was a MPL or r-MPL, one of its rows was selected and the choice implemented; for

---

<sup>8</sup>The complete instructions and the screenshots are presented in Appendix C.

<sup>9</sup>To test for order effects, subjects were randomly assigned to one of the two possible orders of questions. In Appendix B we present the two orders and show that results are not sensitive to them.

TABLE 1: Main Questions

	Question type	Left option	Right option	Values
Q1	MPL	0.5 \$20, 0.5 \$0	\$x	$x \in [0, 20]$
Q2	MPL	\$18	0.5 \$x, 0.5 \$0	$x \in [18, 54]$
Q3	MPL	0.5 \$22, 0.5 \$0	0.5 \$x, 0.5 \$4	$x \in [14, 22]$
Q1r	r-MPL	0.5 \$20, 0.5 \$0	\$x	$x \in [0, 20]$
Q2r	r-MPL	\$18	0.5 \$x, 0.5 \$0	$x \in [18, 54]$
Q3r	r-MPL	0.5 \$22, 0.5 \$0	0.5 \$x, 0.5 \$4	$x \in [14, 22]$
Q4	MPL	\$16	0.8 \$x, 0.2 \$0	$x \in [16, 27]$
Q5	MPL	0.25 \$16, 0.75 \$0	0.2 \$x, 0.8 \$0	$x \in [16, 27]$
Q6	MPL	\$14	0.8 \$x, 0.2 \$0	$x \in [14, 25]$
Q7	MPL	0.25 \$14, 0.75 \$0	0.2 \$x, 0.8 \$0	$x \in [14, 25]$

Notes: We denote by  $p \$x, (1 - p) \$y$  the lottery that pays  $\$x$  with probability  $p$  and  $\$y$  with probability  $(1 - p)$ . In all questions, the Left option stays the same in all rows, while the Right option changes, with values of  $x$  increasing from the top row to the bottom. The last column indicates the range.

r-MPLs, if subjects chose a non-degenerate lottery, each option was given with the specified probability. Second, one control questions (IQ, overconfidence) was randomly selected for payment.

## 4 Results

**Preliminaries.** A total of 165 subjects participated in an experiment run at the University of California, Irvine; all subjects were undergraduate students at that institution. We focus on the 148 subjects who made non-dominated choices in all ten questions.<sup>10</sup>

In discussing behavior in standard MPLs we distinguish between subjects with *monotone* and *non-monotone* choices. The former switch from the Left to the Right option at most once; the latter switch multiple times. With monotone choices, the key measure is the dollar amount linked to the switching point: following standard practice, we use the average dollar amount between the last Left and the first Right choice.

To analyze choices in r-MPLs we use the following measures. We denote by *range91* the range of dollar amounts in which a subject chooses both options with positive probability: the range from the smallest to the largest value for which they indicate numbers between 9 and 1.<sup>11</sup> We say that a subject *exhibits a range91* if numbers other than 0 or 10 appear in more than one row (as choosing it in one row alone could be due to indifference.) As

<sup>10</sup>The remaining 17 subjects are either not paying attention or have non-monotone preferences on money, making the analysis difficult. Including them does not change significantly any of our results; we omit it for brevity.

<sup>11</sup>To be consistent with how we code behavior in standard MPLs, one extreme of *range91* is the average dollar amount between the last row in which the subject chose 10 and the first with values in  $[1, 9]$ ; similarly, the other extreme is the average dollar amount between the last row with values in  $[1, 9]$  and the next.

TABLE 2: Summary Statistics about Ranges

	Q1r (\$20,\$0;50%) vs \$x		Q2r \$18 vs (\$x,\$0;50%)		Q3r (\$22,\$0;50%) vs (\$x,\$4;50%)	
	range91	range64	range91	range64	range91	range64
% of Subjects with non-zero ranges	75% (n = 111)	59% (n = 87)	76% (n = 112)	62% (n = 92)	66% (n = 98)	54% (n = 80)
<b>For non-Zero Ranges</b>						
Av. size in \$ (s.e.)	6.7 (0.35)	3.0 (0.26)	17.2 (0.88)	9.9 (0.88)	4.4 (0.22)	2.5 (0.22)
Av. bottom \$ (s.e.)	5.3 (0.29)	8.1 (0.28)	28.2 (0.59)	31.0 (0.60)	15.1 (0.15)	16.3 (0.19)
Av. top \$ (s.e.)	12.0 (0.23)	11.1 (0.24)	45.4 (0.70)	40.9 (0.76)	19.5 (0.19)	18.8 (0.21)
Av. # rows (s.e.)	9.4 (0.44)	4.9 (0.35)	13.2 (0.56)	8.4 (0.61)	8.9 (0.43)	5.0 (0.44)
Median/ Total # rows	10 / 19	4 / 19	14 / 23	7 / 23	9 / 17	3.5 / 17

Notes: The last five lines report average or median values conditional on exhibiting ranges (standard errors in parenthesis). The last line also includes the total number of rows in each question.

we discussed, any such behavior is not compatible with monotone Expected Utility preferences.

We construct *range64* in a similar way, except that we focus on the range in which both options are chosen with probability at least 40%—indicating 4, 5, or 6. While *range91* captures all values for which the subject prefers to mix, *range64*, obviously a subset, indicates values for which the subject assigns high weight to both options. In the main body, we restrict our attention to these two definitions. Appendix A.2 replicates our analysis also for ranges 8-2 and 5-5, defined similarly, finding coherent results.

#### 4.1 Ranges: Frequency and Size

People *very often* report ranges and these ranges are *big*. Table 2 shows the fraction of subjects who exhibit either kind of range in each question, the dollar width of these ranges, and the number of rows involved.

**Ranges are Frequent.** In all questions, between two-thirds and three-quarter of subjects exhibit *range91*; more than 50% do so for *range64*. Furthermore, the vast majority exhibit ranges in at least one question: only 16% of subjects never exhibit *range91* in any of the three questions, and only 23% never exhibit *range64*. Instead, most exhibit *range91* in multiple occasions: 57% in all three questions, 20% in two, and 7% in one. Similarly, 34% exhibit *range64* in all three questions, 29% in two, and 14% in one. That is: most subjects exhibit ranges in most questions; this is true even if we take restrictive definitions of ranges (like *range64*).

**Ranges are Big.** Conditional of exhibiting a range, ranges are ‘*very large*’: if we look at Q1r, subjects who exhibit range91 (75%) have an average range of size \$6.7. On average, ranges go from \$5.3 to \$12 (median size is \$6.8 and median spans from \$5 to \$12.3). The top of the range is more than twice the bottom. This is a remarkable span, especially because this question measures the certainty equivalent of a lottery that pays \$20 or \$0 with equal chances. Ranges remain ‘*very large*’ also for Q2r and Q3r—the former being much bigger, also reflecting the different scale of payoffs.

Ranges are big also if we look at the number of rows with randomization, instead of the span in dollars. On average, subjects want to randomize in 9.4 rows in Q1r, 13.2 in Q2r, and 8.9 in Q3r (medians are 10, 14, and 9).<sup>12</sup> These are about *half or more* of the rows in each question (there are 19, 23, and 17 rows, respectively). It shows that the large span of range91 is not an artifact of large gaps; rather, subjects consistently choose to randomize in many rows. This should alleviate concerns that ranges are chosen by mistake.

Also range64 is consistently large, with a mean of \$3.0 (median \$2.5) for Q1. The top of the range is 35% above the bottom. We have large ranges also for the other questions. These ranges are big also in terms of number of rows: on average, 4.9, 8.4, and 5.0 rows in Q1r, 2r, and 3r, respectively (medians are 4, 7, 3.5). Thus, there is a sizable span of dollar values, and a sizable number of rows, for which subjects not only want to randomize, but do so giving at least 40% of chance to both options. Overall, subjects are creating very substantial randomizations, with high probabilities for many options.

**Risk Attitude.** We showed that subjects exhibit ranges and that these ranges are big. Where are they located? From Table 2 it is clear that they span substantially *above and below the risk neutral value*. Consider again Q1: the expected value of the lottery is \$10; the average range goes from \$5.3 to \$12. The range includes the risk neutral value in its interior, and extends in both directions. This is not symmetric: ranges extend much deeper into the risk averse direction (lower numbers). But they do span into the risk seeking area: not only the average top of the range is \$12 and the median is \$12.3; of the subjects that exhibit range91 for Q1, 72% have the top part of the range strictly above \$10. Similar results hold for Q2 and Q3, where ratios are 90% and 76%, respectively.<sup>13</sup>

This implies that, conditional on having a range, the majority of our subjects prefers to randomize between a lottery and sure amounts also *above* the expected value of the lottery. We have seen (Section 2) that this is incompatible with risk aversion, even though choices in regular MPLs are typically risk averse (e.g., 55% of choices in Q1-3 are risk averse).

One interpretation is that, while choices in the standard MPLs tend to be risk averse,

---

<sup>12</sup>Moreover, the large majority of these answers are monotone, in the sense that subjects report weakly decreasing numbers within the range as we proceed down the rows. We have monotone answers for 81% of subjects with ranges in Q1r ( $n = 111$ ), 72% in Q2r ( $n = 112$ ), and 82% in Q3r ( $n = 98$ ).

<sup>13</sup>For Q2r the risk seeking part are values below \$36; for Q3r are values above \$18.

r-MPLs allow us to acquire a more nuanced view. If, in line with some of the models above, we interpret ranges as boundaries of the values the agent is sure about—or of the utility functions that are being considered—then our results suggest that agents consider values that fall both in the risk averse *and* in the risk seeking domain. As if they were not fully sure they should be risk averse. When asked to make a precise choice in standard MPLs, however, they tend to fall in the risk averse area. We discuss these choices next.

## 4.2 Ranges and standard MPLs

We use both standard MPLs and r-MPLs for the same three questions. This allows us to ask: Where does the choice in the standard MPL fall with respect to the range expressed in the corresponding r-MPL? In the middle or biased towards one end?

**For subjects without ranges.** We begin with a sanity check. For each question, consider subjects with no range in the r-MPL and with monotonic answers in the MPL. Do answers coincide? This allows us to evaluate if r-MPLs bias responses in a particular way.

We find that answers are highly related and show no particular bias. In all questions, for subjects without a range there is a very high correlation between switching points in MPLs and r-MPLs: correlations are 0.84, 0.85, and 0.72 for Q1, Q2, and Q3 and are significant at 1% level. Moreover, the differences in switching points are distributed around zero, suggesting there is no bias. In other words, subjects who don't express ranges make consistent choices in both formats.<sup>14</sup>

**Defining a measure.** We now turn to subjects with monotone choices in MPLs but also a range in r-MPLs. (We discuss subjects with non-monotone choices below.) To shed light on their behavior, for each subject in this group and relevant question  $Q_i$ , define  $\lambda_{91}^{Q_i} \in \mathbb{R}$  by

$$\text{Standard MPL}^{Q_i} = \lambda_{91}^{Q_i} \cdot \text{Top range91}^{Q_i} + (1 - \lambda_{91}^{Q_i}) \cdot \text{Bottom range91}^{Q_i},$$

where, as names suggest, 'Standard MPL' is the switch point in standard MPLs; 'Top range91' is boundary of range91 in the direction of less risk aversion; 'Bottom range91' is the other.<sup>15</sup> In words,  $\lambda_{91}$  represent where the choice in a standard MPL fall w.r.t. the range expressed in the corresponding r-MPL. Values are in  $[0, 1]$  if and only if the former falls inside the range, and higher (lower) values indicate choices towards more (less) risk-seeking options. Thus,  $\lambda_{91} = 1$  means that the choice in the MPL is exactly at the least risk averse point of range91, while  $\lambda_{91} = 0$  if it is at the most risk averse point. We define  $\lambda_{64}$  analogously.

<sup>14</sup>Moreover, their choices are also not concentrated around salient values. Appendix A.3 presents an in-depth analysis.

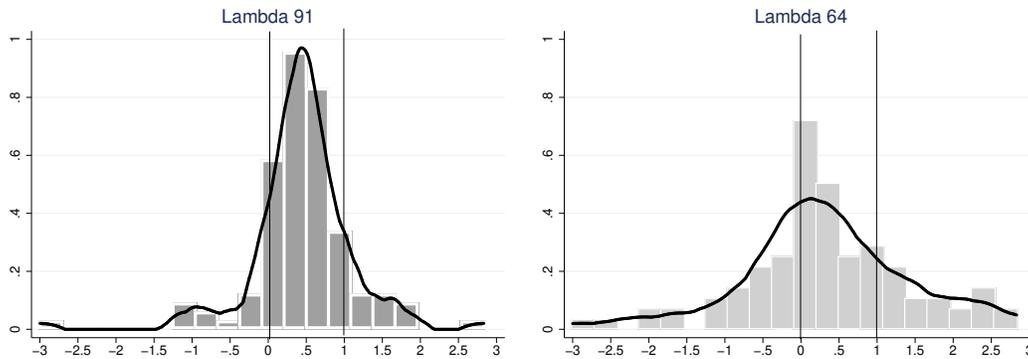
<sup>15</sup>For Q1r and Q3r 'Top range91' and 'Bottom range91' are the highest and lowest numbers in the ranges, respectively. Instead, for Q2r, 'Bottom range91' is the highest number while the 'Top range91' the lowest.

**Distribution of  $\lambda$ s.** Despite some heterogeneity,  $\lambda_{91}$  tends to be in  $[0, 1]$ : this holds for 81% of subjects in Q1, 73% in Q2, and 67% in Q3. This means that: most subjects' choices in standard questions fall within the range identified in r-MPLs. This is not the case, however, for  $\lambda_{64}$ , where the fraction of subjects with values in  $[0, 1]$  are 46%, 51%, and 35%: choices in standard MPLs tend to fall outside this more restrictive range.

Figure 1 plots the distributions of average  $\lambda_{91}$  and  $\lambda_{64}$  across subjects, where we take the average across questions (eliminating 2 outliers for clarity; Figure A.3 in Appendix plots them by question). The picture is clear:  $\lambda_{91}$  tends to be symmetric around 0.5, and concentrated in  $[0, 1]$ , while  $\lambda_{64}$  is far from symmetric and tends to revolve around zero. Indeed, 53% of subjects have  $\lambda_{91}$  between 0.25 and 0.75, while 47% of subjects have  $\lambda_{64}$  below 0.25.

Overall, this suggests that: i) choices in standard MPLs are in the middle of range91; but they are also ii) closer to the *bottom* of range64—the more risk-averse choice.

FIGURE 1: Distribution of Average Values of  $\lambda_{91}$  and  $\lambda_{64}$



Notes: Distribution of  $\lambda_{91}$  (left) and  $\lambda_{64}$  (right), focusing on values between -3 and 3 (this removes 2 outliers). The bottom of the range is the 'more risk-averse' behavior. For both panels, we first estimate  $\lambda$  for each of the question for each subject; then, we average across questions (for subjects with more than one estimated  $\lambda$ ).

We can connect these results with our analysis of risk attitudes and ranges. We have seen that ranges typically extend to the risk seeking domain, but they also tend to be asymmetric, extending further into risk averse areas. If choices in standard MPLs tend to fall in the middle of range91, or at the bottom (more risk averse) of range64, then they will be risk averse. As we suggested above, one possible interpretation is that subjects are unsure of how to evaluate lotteries and contemplate both risk averse and risk seeking options; but when it comes to selecting one option in standard MPLs, they pick a risk averse one, either being cautious, or because the range they consider is asymmetric.

TABLE 3: Ranges, Individual Characteristics, and non-Monotone Behavior

	Range 91			Range 64		
	Ind. Range	Range Freq.	Range Size	Ind. Range	Range Freq.	Range Size
Risk Q1	0.07 (0.44)	<b>0.16*</b> (0.07)	0.09 (0.34)	0.13 (0.15)	<b>0.20**</b> (0.03)	0.07 (0.41)
Risk Q2	0.13 (0.17)	<b>0.19**</b> (0.03)	<b>0.24***</b> (0.01)	0.05 (0.55)	0.15 (0.11)	0.14 (0.14)
Risk Q3	0.01 (0.90)	0.05 (0.56)	-0.01 (0.93)	0.08 (0.38)	0.12 (0.16)	-0.04 (0.67)
C-bias	-0.13 (0.18)	<b>-0.16*</b> (0.10)	<b>-0.20**</b> (0.04)	<b>-0.18*</b> (0.06)	<b>-0.26**</b> (0.01)	<b>-0.26***</b> (0.01)
Non-Mon. Q1-7	<b>0.24**</b> (0.00)	<b>0.21**</b> (0.01)	<b>0.36**</b> (0.00)	<b>0.23**</b> (0.01)	<b>0.26**</b> (0.00)	<b>0.50**</b> (0.00)

Notes: Pearson pairwise correlations with significance level in parenthesis; \*\*\*, \*\*, and \* indicates significance at 1%, 5%, and 10% level, respectively. For C-Bias we have  $n = 108$  observations (the number of subjects that report monotonic choices in Q4-Q7).

### 4.3 Ranges and Individual Characteristics

How do ranges relate to individual characteristics such as risk attitudes, Certainty Bias, IQ, and overconfidence?

**Measures.** The answers to standard MPLs in Q1-2-3 give us three continuous measures of risk aversion: the certainty equivalent of a lottery (Q1), the lottery equivalent of a sure amount (Q2), and the lottery equivalent of a lottery (Q3). Following Chapman et al. (2019b), we treat these measures separately. For ease of comparisons, values are normalized such that higher numbers mean more risk aversion.

The difference between the answers to Q4 and Q5, and to Q6 and Q7, give us two measures of the common ratio effect, or Certainty Bias (Allais, 1953). We take the weighted average (dividing by the fixed amount on the MPLs) and denote it *C-bias*.<sup>16</sup> Higher numbers denote higher certainty bias.<sup>17</sup>

**Results.** Table 3 shows the correlation between these measures with: an indicator on whether the subject exhibits the range at all (Ind. Range); the number of questions with ranges (Range Freq.); and the average dollar size of ranges (Range Size).<sup>18</sup> Results are clear.

First, there are some but limited relations between risk attitudes and ranges. Risk Q2 is significantly related to range91, with more risk-averse subjects reporting ranges for more questions and larger ranges. But this does not extend robustly to other measures of risk,

<sup>16</sup>That is, the variable C-Bias is constructed by  $\frac{1}{16}(Q4 - Q5) + \frac{1}{14}(Q6 - Q7)$ , where Q4 is value in Q4, etc.

<sup>17</sup>For subjects for whom we have this measure (which requires monotone answers), 58%, 6%, and 35% have a positive, zero, and negative values. If we allow a small band around zero (between -0.1 and 0.1), values become 44%, 27%, and 29%, respectively.

<sup>18</sup>This is the weighted average of range sizes, normalized by the expected value of the Left Option. In Appendix A.5 we show that the results remain the same if we normalize by the maximum possible size of the range in each question (last column of Table 1).

and loses significance for range64.<sup>19</sup>

Second, Certainty Bias relates negatively to the frequency and size of ranges. Subjects with more such bias have less frequent and smaller ranges. There is thus a connection between these two forms of violations of Expected Utility.

**IQ and Overconfidence.** We also analyze the relation between ranges and IQ and measures of overconfidence (overestimation, overplacement, and overprecision). We find no robust evidence of any relation; details are discussed in Appendix A.6.

#### 4.4 Ranges and Non-monotone choices

We conclude with a connection with non-monotone choices. Experiments that use standard MPLs usually find a sizable fraction of non-monotone answers, i.e., multiple switches between the left and right option. This is true in our sample as well: 18% of subjects display non-monotonic choices at least once in Q1, Q2, or Q3. The typical approach is to disregard these choices, treating them as solely noise. But what if they are instead informative? For example, subjects may be randomizing in each line and thus exhibit non-monotone behavior—as already suggested by Chew et al. (2019).

Our first observation is that non-monotone behavior is very strongly related with ranges. This appears in the last row of Table 3, showing that violations of monotonicity very strongly correlate with i) exhibiting ranges, ii) doing so more often, and iii) having larger ranges. Note how highly significant all these results are.

A second approach is to classify subjects into types based on 1) choice monotonicity in MPLs, and 2) existence of ranges<sup>91</sup> in r-MPLs. This leaves us with four types:

- 1. Monotone and no ranges.** These are the types predicted by standard theory. In our sample, only 15% (22 subjects) are monotone in all MPLs and never report ranges.
- 2. Monotone and ranges.** About a half of our subjects (51% or 75 subjects) belong to this category.
- 3. Non-monotone and ranges.** A third of subjects (33% or 49 subjects) display non-monotone behavior in MPLs and exhibit ranges r-MPLs.
- 4. Non-monotone and no ranges.** This category is essentially non-existent in our sample. We observe only 2 subjects (1%) of this type.

This classification shows two key aspects of our data. First, ‘standard’ subjects—monotone and no ranges—are a real minority in our sample. The majority has ranges,

---

<sup>19</sup>The fact that risk measures from Q1 and Q2 relate differently is not surprising in light of Chapman et al. (2019a,b), where only the latter is related to a number of demographics and other characteristics.

and tends to have monotone choices. Importantly, of the subjects that ever violate monotonicity, essentially all of them also exhibit ranges, showing a connection between these tendencies.

## 5 Discussion

This paper introduces a simple method—a modified Multiple-Price-List—to capture the range of values for which subjects want to randomize between two options. In an experiment, we find that decision-makers express this desire 1) very frequently and 2) for very wide ranges of values, that 3) spill into the risk-seeking realm; moreover, 4) choices in standard questions tend to fall either in the middle or in the bottom of such ranges and, 5) ranges are related to Certainty Bias and non-Monotonic choices in standard questions.

Overall, our results show that preferences for randomization, and the corresponding violations of Expected Utility, not only exist, but are in fact prevalent for wide ranges encompassing values typically used in economic problems. These phenomena are thus far from being restricted to smartly-constructed, knife-edge cases; instead, they apply broadly. Ignoring them may lead to the wrong choices in modeling or in deriving conclusions from existing data, e.g., about welfare.

More broadly, one may view our results as contributing to the growing body of work that tries to measure difficulty in making comparisons—incompleteness, imprecision, cognitive uncertainty; as we have seen, preferences for randomization may be linked to incomplete preferences with the idea that agents may randomize when unsure. One of the goals of this literature is to connect difficulty in comparisons with existing behavioral phenomena. A large theoretical literature has linked incomplete preferences with status quo bias, endowment effect, certainty bias, ambiguity aversion, stochastic choice, time preferences, the attraction effect, and other aspects;<sup>20</sup> a theoretical and empirical literature has linked preference imprecision or cognitive uncertainty to these and other phenomena;<sup>21</sup> a separate empirical literature has linked stochastic choice to desire to randomize.<sup>22</sup> Here we show that the desire to randomize, related to incompleteness, is very prevalent for large spans of value.

---

<sup>20</sup>Bewley (1986); Ghirardato et al. (2004); Masatlioglu and Ok (2005); Ok and Masatlioglu (2007); Gilboa et al. (2010); Ortoleva (2010); Masatlioglu and Ok (2014); Cerreia-Vioglio et al. (2015); Ok et al. (2015); Cerreia-Vioglio et al. (2019a); Ok and Nishimura (2019) and many references therein.

<sup>21</sup>Butler and Loomes (2007, 2011); Khaw et al. (2018); Enke and Graeber (2019); Gabaix (2019).

<sup>22</sup>Agranov and Ortoleva (2017) and references therein.

## References

- AGRANOV, M., P. J. HEALY, AND K. NIELSEN (2020): “Non-Random Randomization,” Mimeo.
- AGRANOV, M. AND P. ORTOLEVA (2017): “Stochastic Choice and Preferences for Randomization,” *Journal of Political Economy*, 125, 40–68.
- ALLAIS, M. (1953): “Le comportement de l’homme rationnel devant le risque: Critique des postulats et axiomes de l’école Américaine,” *Econometrica*, 21, 503–546.
- BECKER, G. M., M. H. DEGROOT, AND J. MARSCHAK (1964): “Measuring utility by a single-response sequential method,” *Behavioral Science*, 9, 226–232.
- BEWLEY, T. F. (1986): “Knightian Decision Theory: Part I,” *Cowles Foundation discussion paper*, 807.
- BUTLER, D. AND G. LOOMES (2007): “Imprecision as an account of the preference reversal phenomenon,” *The American Economic Review*, 277–297.
- (2011): “Imprecision as an account of violations of independence and betweenness,” *Journal of Economic Behavior & Organization*, 80, 511–522.
- CERREIA-VIOGLIO, S. (2010): “Maxmin Expected Utility on a Subjective State Space: Convex Preferences under Risk,” Mimeo Bocconi University.
- CERREIA-VIOGLIO, S., D. DILLENBERGER, AND P. ORTOLEVA (2015): “Cautious Expected Utility and the Certainty Effect,” *Econometrica*, 83, 693–728.
- (2019a): “An Explicit Representation for Disappointment Aversion and Other Betweenness Preferences,” *Theoretical Economics (forthcoming)*.
- CERREIA-VIOGLIO, S., D. DILLENBERGER, P. ORTOLEVA, AND G. RIELLA (2019b): “Deliberately Stochastic,” *American Economic Review*, 109, 2425–45.
- CETTOLIN, E. AND A. RIEDL (2019): “Revealed preferences under uncertainty: Incomplete preferences and preferences for randomization,” *Journal of Economic Theory*, 181, 547–585.
- CHAPMAN, J., M. DEAN, P. ORTOLEVA, E. SNOWBERG, AND C. CAMERER (2019a): “Econographics,” Mimeo Princeton University.
- (2019b): “Willingness-To-Pay and Willingness-To-Accept are Probably Less Correlated than You Think,” Mimeo, Princeton University.
- CHEW, S. H., B. MIAO, Q. SHEN, AND S. ZHONG (2019): “Multiple-Switching Behavior in Choice-List Elicitation of Risk Preference,” Mimeo, National University of Singapore.
- COHEN, M., J.-Y. JAFFRAY, AND T. SAID (1985): “Individual behavior under risk and under uncertainty: An experimental study,” *Theory and Decision*, 18, 203–228.
- (1987): “Experimental comparison of individual behavior under risk and under uncertainty for gains and for losses,” *Organizational Behavior and Human Decision Processes*, 39, 1–22.
- CONDON, D. M. AND W. REVELLE (2014): “The International Cognitive Ability Resource: Development and initial validation of a public-domain measure,” *Intelligence*, 43, 52–64.

- CUBITT, R., D. NAVARRO-MARTINEZ, AND C. STARMER (2015): "On preference imprecision," *Journal of Risk and Uncertainty*, 50, 1–34.
- DEKEL, E. (1986): "An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom," *Journal of Economic Theory*, 40, 304–318.
- DUBOURG, W. R., M. W. JONES-LEE, AND G. LOOMES (1994): "Imprecise preferences and the WTP-WTA disparity," *Journal of Risk and Uncertainty*, 9, 115–133.
- (1997): "Imprecise preferences and survey design in contingent valuation," *Economica*, 681–702.
- DUBRA, J., F. MACCHERONI, AND E. A. OK (2004): "Expected utility theory without the completeness axiom," *Journal of Economic Theory*, 115, 118–133.
- DWENGER, N., D. KÜBLER, AND G. WEIZSÄCKER (2018): "Flipping a Coin: Evidence from Laboratory and Field," *Journal of Public Economics*, 167, 240–250.
- ENKE, B. AND T. GRAEBER (2019): "Cognitive Uncertainty," Mimeo, Harvard University.
- FELDMAN, P. AND J. REHBECK (2020): "Revealing a preference for mixing: An experimental study of risk," Mimeo.
- GABAIX, X. (2019): "Behavioral inattention," in *Handbook of Behavioral Economics: Foundations and Applications*, ed. by D. Bernheim, S. DellaVigna, and D. Laibson, Elsevier, vol. 2, 261–343.
- GHIRARDATO, P., F. MACCHERONI, AND M. MARINACCI (2004): "Differentiating ambiguity and ambiguity attitude," *Journal of Economic Theory*, 118, 133–173.
- GILBOA, I., D. SCHMEIDLER, F. MACCHERONI, AND M. MARINACCI (2010): "Objective and subjective rationality in a multi-prior model," *Econometrica*, 78, 755–770.
- GNEEZY, U. AND J. POTTERS (1997): "An experiment on risk taking and evaluation periods," *The Quarterly Journal of Economics*, 112, 631–645.
- KAHNEMAN, D., J. KNETSCH, AND R. THALER (1991): "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias," *Journal of Economic Perspectives*, 5, 193–206.
- KHAW, M. W., Z. LI, AND M. WOODFORD (2017): "Risk aversion as a perceptual bias," NBER Working Paper No. 23294.
- (2018): "Cognitive imprecision and small-stakes risk aversion," NBER Working Paper No. 24978.
- KIRCHER, P., S. LUDWIG, AND A. SANDRONI (2013): "On the Difference between Social and Private Goods," *The BE Journal of Theoretical Economics*, 13, 1–27.
- MACHINA, M. J. (1985): "Stochastic choice functions generated from deterministic preferences over lotteries," *The Economic Journal*, 95, 575–594.
- MASATLIOGLU, Y. AND E. A. OK (2005): "Rational Choice with status quo bias," *Journal of Economic Theory*, 121, 1–29.
- (2014): "A Canonical Choice Model with Initial Endowment," *Review of Economic Studies*, 81, 851–883.

- MIAO, B. AND S. ZHONG (2018): "Probabilistic social preference: how Machina's Mom randomizes her choice," *Economic Theory*, 65, 1–24.
- OK, E. AND H. NISHIMURA (2019): "Preference Structures," Mimeo, New York University.
- OK, E., P. ORTOLEVA, AND G. RIELLA (2012): "Incomplete Preferences under Uncertainty: Indecisiveness in Beliefs vs. Tastes," *Econometrica*, 80, 1791–1808.
- (2015): "Revealed (P)Reference Theory," *American Economic Review*, 105, 299–321.
- OK, E. A. AND Y. MASATLIOGLU (2007): "A theory of (relative) discounting," *Journal of Economic Theory*, 137, 214–245.
- ORTOLEVA, P. (2010): "Status Quo Bias, Multiple Priors and Uncertainty Aversion," *Games and Economic Behavior*, 69, 411–424.
- ORTOLEVA, P. AND E. SNOWBERG (2015): "Overconfidence in Political Behavior," *American Economic Review*, 105, 504–535.
- PERMANA, Y. (2020): "Why do people prefer randomisation? An experimental investigation," *Theory and Decision*, 88, 73–96.
- QUIGGIN, J. (1982): "A theory of anticipated utility," *Journal of Economic Behavior & Organization*, 3, 323–343.
- RUBINSTEIN, A. (2002): "Irrational diversification in multiple decision problems," *European Economic Review*, 46, 1369–1378.
- SOPHER, B. AND J. M. NARRAMORE (2000): "Stochastic choice and consistency in decision making under risk: An experimental study," *Theory and Decision*, 48, 323–350.

# Online Appendix

## A Additional Analysis

### A.1 Steps in MPLs and r-MPLs

The list below lists the steps used in each of our MPL and r-MPL questions (all amounts are in US dollars):

- Q1 and Q1r: 0,4,6,7,7.5,8,8.5,9,9.5,10,10.5,11,11.5,12,12.5,13,14,16,20
- Q2 and Q2r: 18,21,24,27,30,33,33.50,34,34.50,35,35.50,36,36.50,37,37.50,38,38.50,39,42,45,48,51,54
- Q3 and Q3r: 14,14.50,15,15.50,16,16.50,17,17.50,18,18.50,19,19.50,20,20.50,21,21.50,22
- Q4 and Q5: 16,18,18.50,19,19.50,20,20.50,21,21.50,22,22.50,23,23.50,24,24.50,25,27
- Q6 and Q7: 14,16,16.50,17,17.50,18,18.50,19,19.50,20,20.50,21,21.50,22,22.50,23,25

### A.2 Other Types of Ranges

In the main body of the paper we focused on ranges ‘9-1’ and ‘6-4.’ We now report key measures for range ‘8-2’ and ‘5-5,’ defined analogously.

TABLE A.1: Summary Statistics about Other Types of Ranges

	Q1r (\$20,\$0;50%) vs \$x		Q2r \$18 vs (\$x,\$0;50%)		Q3r (\$22,\$0;50%) vs (\$x,\$4;50%)	
	range82	range55	range82	range55	range82	range55
% of Subjects with non-zero ranges	70% (n = 103)	43% (n = 64)	73% (n = 108)	49% (n = 72)	61% (n = 91)	33% (n = 49)
<b>For non-Zero Ranges</b>						
Av. size in \$ (s.e.)	5.2 (0.30)	2.9 (0.35)	14.2 (0.84)	9.7 (1.08)	3.6 (0.22)	1.6 (0.24)
Av. bottom \$ (s.e.)	6.5 (0.29)	8.2 (0.37)	28.2 (0.58)	30.9 (0.69)	15.4 (0.15)	16.3 (0.26)
Av. top \$ (s.e.)	11.7 (0.23)	11.1 (0.28)	45.5 (0.69)	40.6 (0.90)	19.3 (0.19)	18.9 (0.27)
Av. # rows (s.e.)	7.9 (0.40)	4.5 (0.46)	11.7 (0.56)	8.0 (0.76)	7.7 (0.41)	5.3 (0.65)
Median/ Total # rows	7 / 19	3 / 19	13 / 23	5 / 23	8 / 17	4 / 17

**Notes:** The last five lines report average or median values conditional on exhibiting ranges (standard errors in parenthesis). The last line also includes the total number of rows in each question.

We find that 18% of subjects never report range82 and 29% never report range55. The remaining subjects report these ranges at least once. Specifically, 49% (14%) of subjects report range82 (range55) in all three questions, 23% (26%) do so for two out of three questions, and 10% (31%) do so in one question.

### A.3 Subjects with No Ranges and Monotone Behavior

FIGURE A.1: Difference between choices in  $Q_x$  and  $Q_{xr}$  for subjects who are monotone in  $Q_x$  and do not exhibit a range in  $Q_{xr}$

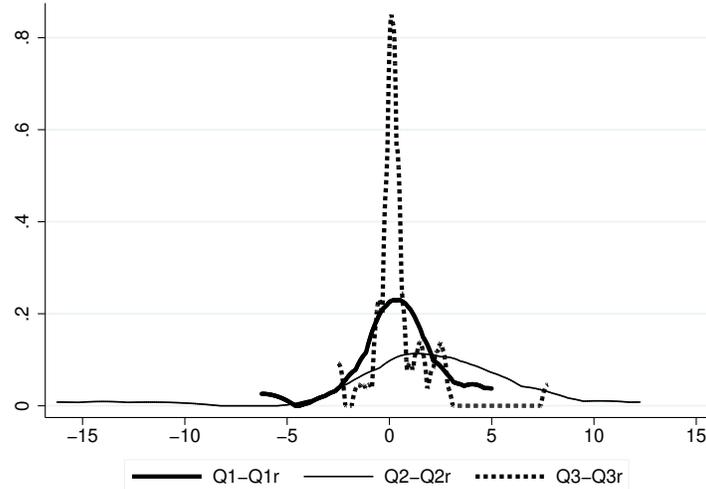
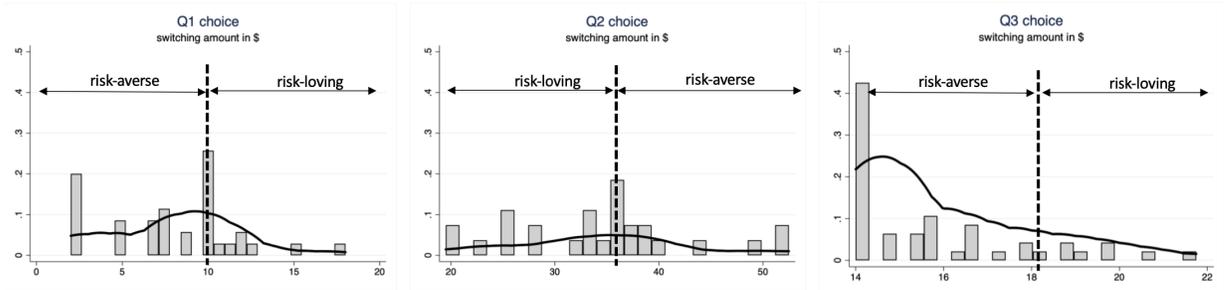


FIGURE A.2: Switching point in regular MPLs for Monotone Subjects who do not exhibit ranges in that question



We expand our analysis on the comparison between choices in standard MPLs and range MPLs for subjects who did not exhibit ranges and are monotone in the standard MPL. Figure A.1 shows the kernel distributions of the differences between switching points in the two formats, for each question. The distributions appear quite symmetric around zero; the value zero is in the 95% confidence interval of the estimated mean (for each question separately). For Q1 and Q2 we cannot reject the null that this difference is equal to zero at the standard 5% level ( $p = 0.38$  and  $p = 0.17$ , respectively, according to two-sided T-test). For Q3-Q3r we marginally reject this null ( $p = 0.05$ ); however, if we remove one

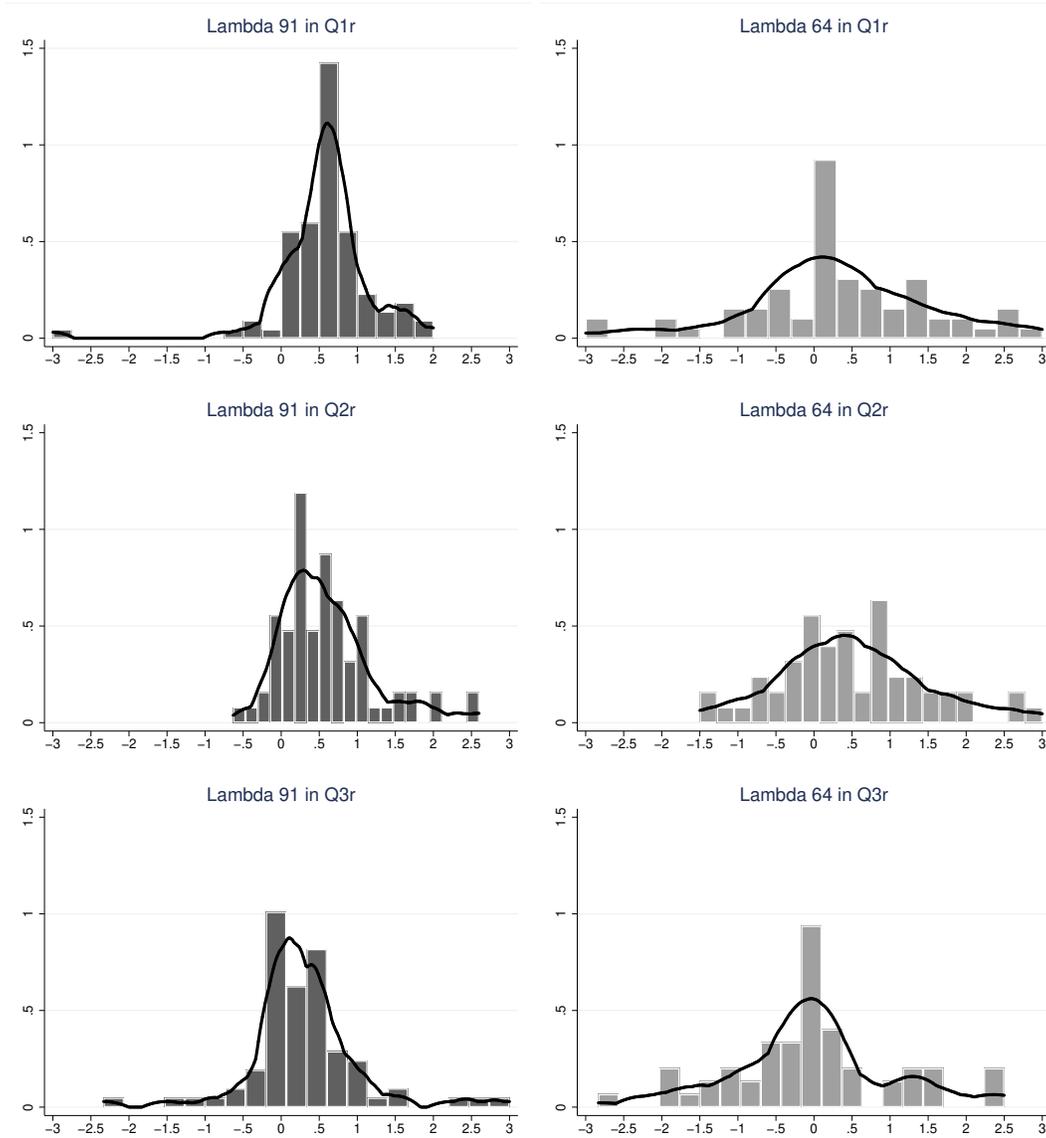
outlier observation, we obtain  $p > 0.05$  for Q3 as well.

Figure A.2 shows the distribution of answers in regular MPLs for questions where subjects who do not exhibit ranges. The key observation is that these answers are not concentrated around salient values (e.g., the expected value) but rather are spread throughout.

## A.4 Individual Estimates of Lambda

Figure A.3 depicts the distributions of  $\lambda_{91}$  and  $\lambda_{64}$  in each question.

FIGURE A.3: Distribution of  $\lambda_{91}$  and  $\lambda_{64}$  for Q1r, Q2r, and Q3r



## A.5 Alternative Measure of Range Size and Individual Characteristics

Table A.2 repeats the relevant part of Table 3 for an alternative definition of range size, where the normalization is made using the maximum possible range size.

TABLE A.2: Size of Ranges, Individual Characteristics, and non-Monotone Behavior

	Size of Range 91	Size of Range 64
Risk Q1	0.12 (0.20)	0.09 (0.34)
Risk Q2	<b>0.26***</b> (0.004)	<b>0.16*</b> (0.09)
Risk Q3	-0.01 (0.89)	-0.04 (0.63)
C-bias	<b>-0.19**</b> (0.04)	<b>-0.21**</b> (0.03)
Non-Mon. Q1-7	<b>0.33***</b> (0.00)	<b>0.51***</b> (0.00)

Notes: Pearson pairwise correlations with significance level in parenthesis; \*\*\*, \*\*, and \* indicates significance at 1%, 5%, and 10% level, respectively. For C-Bias we have  $n = 108$  observations (the number of subjects that report monotonic choices in Q4-Q7).

## A.6 Relation between Ranges, IQ and Overconfidence

We now explore more in detail the relation between the tendency to exhibit ranges and measures of IQ and overconfidence. Table A.3 presents pairwise correlations.

We have two different measures of IQ: six matrices from the ICAR database and the three CRT questions. The total IQ is the average score across the two, measured by  $\frac{\#correct\ ICAR}{6} + \frac{\#correct\ CRT}{3}$ .

The overconfidence measures reported in the last four rows are computed following standard practice. Overestimation is the difference between how many ICAR questions a subject thinks she solved correctly minus how many she actually solved correctly. Overplacement is the reported rank minus actual rank in a sample of the 100 randomly selected adults in the US (obtained from Chapman et al. 2019b). Overprecision 1 and Overprecision 2 are calculated based on the answers subjects give to the trivia question ("when was the land phone invented?") and the confidence that they have in their answer being correct. We follow the approach in Ortoleva and Snowberg (2015): first, we construct a measure of accuracy by taking the absolute value between the reported year and the actual year; then, we run a regression in which we try to predict confidence with a 4th degree polynomial of accuracy. The residual is our measure of overprecision. Overconfidence 1 uses the confidence measure constructed from the qualitative question "how confident you are in your answer?" (admitting four possible answers), while overconfidence 2 uses the confidence measure constructed from the question "what is the probability that you answered correctly?".

Table A.3 shows that there is no systematic relationship between the IQ and the over-

TABLE A.3: Relation between Ranges, IQ and Overconfidence

	Range 91			Range 64		
	Ind. Range	Range Freq.	Range Size	Ind. Range	Range Freq.	Range Size
# correct ICAR	0.05 (0.54)	0.14* (0.10)	0.04 (0.61)	0.06 (0.46)	0.07 (0.37)	-0.05 (0.55)
# correct CRT questions	-0.07 (0.39)	-0.11 (0.18)	-0.24*** (0.00)	-0.09 (0.28)	-0.17** (0.04)	-0.26*** (0.001)
total IQ	-0.001 (0.99)	0.04 (0.64)	-0.10 (0.24)	-0.004 (0.96)	-0.04 (0.66)	-0.17** (0.04)
Overestimation	-0.02 (0.77)	-0.13 (0.11)	-0.12 (0.16)	0.02 (0.77)	-0.06 (0.50)	-0.03 (0.75)
Overplacement	0.003 (0.97)	0.11 (0.17)	0.15* (0.06)	-0.01 (0.91)	0.08 (0.36)	0.05 (0.58)
Overprecision 1	0.14 (0.21)	0.23** (0.04)	0.08 (0.47)	0.16 (0.14)	0.16 (0.14)	-0.03 (0.75)
Overprecision 2	0.15* (0.07)	0.12 (0.15)	0.12 (0.15)	0.15* (0.06)	0.12 (0.15)	0.07 (0.38)

Notes: Pearson pairwise correlations are reported alongside with p-values: \*\*\*, \*\*, and \* indicate significance at 1%, 5%, and 10%, respectively. The average size of the range is computed as the weighted average of the size of the range in dollars weighted by the expected value of the left option in the question.

confidence and the tendency to exhibit ranges.

### A.7 Relation between Ranges and Questionnaire Answers

At the end of the experiment, we asked subjects the following question. *“In one of the Parts of the experiment you were asked to specify a number between 0 and 10 that determined the probability of receiving the left or the right option. Did you ever choose a number that was different from 0 or 10? If so, can you tell us why? Please elaborate if you can.”* In our main sample of subjects, 62% indicated that they used numbers others than 0 and 10 and gave more or less elaborate reasons for doing so; 25% said that they did not use numbers others than 0 or 10; and 13% did not respond to this question.

Despite the fact that the questionnaire is not incentivized, we find that answers written by subjects are meaningful, consistent with their choices in the experiment, and informative about the mechanism underlying their decision. In particular, *all* subjects that reported that they randomized between the Left and the Right options in some of the questions indeed did so. Table A.4 shows that there is an extremely strong and significant correlation between the answers subjects provide in the questionnaire and their actual behavior in the experiment. In particular, subjects who report the use of ranges are very likely to use them, are more likely to report ranges for a higher number of questions, and have higher average range sizes. This holds true irrespectively of which range measure we use (range91 or range64).

What are the reasons that subjects provide for using ranges? Many subjects indicate that they were not sure whether the Left or the Right option is better for them and preferred to choose both. Here are a few examples:

- “yes because I wanted to have a chance at both options”
- “Yes, because I was not sure what to pick.”

TABLE A.4: Relation between Ranges and Answers in the Questionnaire

	Ind. Range	Range 91 Range Freq.	Range Size	Ind. Range	Range 64 Range Freq.	Range Size
Answer 'yes' in the questionnaire	0.72*** (0.00)	0.72*** (0.00)	0.50*** (0.00)	0.60*** (0.00)	0.56*** (0.00)	0.33*** (0.00)

Notes: Pearson pairwise correlations are reported alongside with p-values: \*\*\*, \*\*, and \* indicate significance at 1%, 5%, and 10%, respectively. The average size of the range is computed as the weighted average of the size of the range in dollars weighted by the expected value of the left option in the question.

- “yes, I was unsure of my decision so I decided to let the computer choose”
- “yes because I was unsure of which would be the best decision and so I decided to just leave it up to probability.”
- “Yes, I was not sure I wanted to take the risk.”
- “Yes, I wasn’t sure which answer would be best, so i decided to let the computer decide for me.”

At the same time subjects that tend not to use ranges explicitly explain that they prefer to make the choice themselves. Here are a few examples from the ‘no’ category:

- “No, because either the left option was better than the right or vice versa.”
- “I always selected 0 or 10 because it is much better to be certain than to leave it up to possibilities.”
- “I didn’t choose a number different from 0 or 10 because I figured I can make the decision myself ”
- “no because it was obvious that was the better choice ”

Overall, we find that subjects’ provide a variety of reasons for reporting ranges, many of which are reminiscent of the mechanisms described by Non-Expected Utility frameworks. Importantly, subjects’ answers are consistent with their choices in the incentivized part of the experiment, which suggests that subjects are aware of their preferences and make these choices consciously.

## B Structure of the Experiment and Order Effects

To investigate the possibility of order effects, we used two different orders across subjects with randomization at a session level; moreover, some parts had questions ordered randomly (at an individual level). Table B.5 illustrates the structure.

TABLE B.5: Two Orders of Questions

	Order A		Order b	
	questions	order	questions	order
Part I	Q1, Q2, Q3	random	Q6, Q7, Q4, Q5	fixed
Part II	Q4, Q5, Q6, Q7	fixed	Q1r, Q2r, Q3r	random
Part III	Risk1 and Risk2	random	Risk1 and Risk2	random
Part IV	Q1r, Q2r, Q3r	random	Q1, Q2, Q3	random
Part V	IQ + overconfidence	fixed	IQ + overconfidence	fixed
Part VI	Questionnaire		Questionnaire	

Notes: Random order indicates that the order of questions in this part of the experiment was randomized across subjects. Otherwise, fixed order was implemented for all subjects.

We find no significant differences in subjects' behavior depending on the order of questions they encountered. Table B.6 replicates our key summary statistics of ranges (similar to Table 2 in the main body of the paper), for Order A and Order B separately.

Results are broadly consistent. For example, among subjects facing Order A we have 11% of subjects who never report range91 and 18% who never report range64. The remaining subjects report ranges at least once. Specifically, 64% (45%) of subjects report range91 (range64) in all three questions, 18% (26%) do so for two out of three questions, and 7% (11%) do so in one question. The distribution of types is similar for subjects facing Order B. Specifically, we have 21% of subjects who never report range91 and 28% who never report range64. Among the remaining subjects, 49% (24%) of subjects report range91 (range64) in all three questions, 21% (32%) do so for two out of three questions, and 8% (16%) do so in one question.

TABLE B.6: Summary Statistics about Ranges of Subjects in Two Orders

	Q1r (\$20,\$0;50%) vs \$x		Q2r \$18 vs (\$x,\$0;50%)		Q3r (\$22,\$0;50%) vs (\$x,\$4;50%)	
	range91	range64	range91	range64	range91	range64
<b>ORDER A (73 subjects)</b>						
% of Subjects with non-zero ranges	82% (n = 60)	70% (n = 51)	84% (n = 61)	71% (n = 52)	70% (n = 51)	58% (n = 42)
<b>For non-Zero Ranges</b>						
Av. size in \$ (s.e.)	6.85 (0.47)	2.90 (0.29)	17.03 (1.19)	9.26 (1.18)	4.84 (0.28)	2.73 (0.33)
Av. bottom \$ (s.e.)	5.08 (0.40)	8.08 (0.34)	28.23 (0.74)	31.47 (0.73)	15.18 (0.20)	16.60 (0.27)
Av. top \$ (s.e.)	11.92 (0.33)	10.98 (0.33)	45.26 (0.97)	40.73 (0.94)	20.02 (0.22)	19.33 (0.27)
Av. # rows (s.e.)	9.5 (0.61)	4.8 (0.42)	13.2 (0.74)	8.4 (0.78)	9.7 (0.57)	5.5 (0.66)
Median/Total # rows	9/19	5/19	14/23	7/23	11/17	4/17
<b>ORDER B (75 subjects)</b>						
% of Subjects with non-zero ranges	68% (n = 51)	48% (n = 36)	68% (n = 51)	49% (n = 40)	53% (n = 47)	51% (n = 38)
<b>For non-Zero Ranges</b>						
Av. size in \$ (s.e.)	6.50 (0.51)	3.11 (0.47)	17.32 (1.32)	10.63 (1.34)	3.89 (0.31)	2.17 (0.28)
Av. bottom \$ (s.e.)	5.58 (0.44)	8.10 (0.47)	28.23 (0.95)	30.46 (1.00)	15.01 (0.21)	15.99 (0.25)
Av. top \$ (s.e.)	12.09 (0.32)	11.22 (0.34)	45.54 (1.03)	41.09 (1.27)	18.90 (0.30)	18.16 (0.30)
Av. # rows (s.e.)	9.3 (0.65)	4.9 (0.61)	13.2 (0.86)	8.5 (0.99)	7.8 (0.62)	4.4 (0.55)
Median/Total # rows	11/19	4/19	15/23	6.5/23	8/17	3/17

Notes: The last five lines for each order report average or median values conditional on exhibiting ranges (standard errors in parenthesis). The last line for each order also includes the total number of rows in each question.

## C Instructions

**General Instructions.** Welcome! This is an experiment designed to study decision-making. The instructions are simple, and if you follow them you may earn a considerable amount of money.

Please **turn off your cell phones** and **do not use them during the experiment**. Please do not talk with others. Also, please do not open any other applications or internet windows on the computer.

**Structure of the Experiment.** The main section of the experiment consists of **4 parts** with a total of **12 questions**. Once you are finished, we will ask you a few additional short questions. The experiment is thus very short. Please think carefully about each choice.

At the end of each Part, the computer will tell you to wait to proceed until prompted: [please do so](#).

Let us highlight from the start that in the main part of the experiment **there are no right or wrong answers**. We are only interested in studying **your preferences**.

**Lotteries.** In many questions, we will ask you to choose between lotteries. Here is an example of a lottery:

50% chance of \$10  
50% chance of \$5

This lottery pays either \$10, with probability 50%, or \$5, with probability 50%. To determine which, the computer will randomly draw an (integer) number between 1 and 100, where each number is equally likely to be drawn. If the drawn number is less or equal to 50, the lottery will pay \$10. If the drawn number is above 50, the lottery will pay \$5. Thus, it pays either \$5 or \$10 with equal probability.

Depending on the questions, the probabilities involved could be different: for example, they could be 25%, 75%, etc. In some cases, the lottery will involve no chance at all: for example, the option may just pay \$12. In all cases, the outcome of lotteries will be determined by the computer using the probabilities specified.

**Your Payment.** Your payment consists of three components:

- First, the computer randomly chooses one of the 12 questions from the main part of the experiment. Each question is equally likely to be selected. Some questions will have several rows, in each of which you will be prompted to make a choice. If the selected question has more than one row in it, then computer also randomly chooses one of the rows in the selected question. Each row is equally likely to be selected.

Your choice in the selected row of the selected question will be the first component of your final payment in this experiment.

- Second, you will receive additional payment for short questions that you will answer at the end of the experiment (after completing the main part of the experiment). You will see the exact instructions on how the short tasks will be paid on your screen.
- Third, you will receive \$12 for showing up and completing the experiment

**PART I.** There are 3 questions in this part. Each question consists of several rows. In each row, there are two options: the Left Option and the Right Option. Here is an example of a question with 5 rows:

50% chance of \$8 50% chance of \$5	<input checked="" type="radio"/> <input type="radio"/>	\$5
50% chance of \$8 50% chance of \$5	<input type="radio"/> <input type="radio"/>	\$6
50% chance of \$8 50% chance of \$5	<input type="radio"/> <input type="radio"/>	\$6.50
50% chance of \$8 50% chance of \$5	<input type="radio"/> <input type="radio"/>	\$7
50% chance of \$8 50% chance of \$5	<input type="radio"/> <input checked="" type="radio"/>	\$8

**For each row**, you must **select one of the two available options**: the Left or the Right one.

Note: **the option on the Left is always the same**. The option on the Right instead changes: it **pays more money as we go down the rows**. This will be the case in **all** questions.

Also, note that in some of the rows, one of the two options pays as much, or more, than the other. This is the case in the first and in the last rows above:

- In the **first row**, the Left option pays either \$8 or \$5, while the Right option pays \$5 for sure. Therefore, the Left Option pays at least as much as the Right Option.
- In the **last row**, the Left option pays either \$8 or \$5, while the Right option pays \$8 for sure. Therefore, the Right Option pays at least as much as the Left Option.

In cases like these, the option that yields higher payoff will be **preselected** for you (indicated by the filled orange circle). You can change this if you wish.

Recall that each question is equally likely to be selected for payment, and that each row within a question is equally likely to be selected for payment.

Also recall that there are **no right or wrong answers**: we are only interested in studying **your preferences**. Finally, there are only **3 questions** in this part, so please think carefully about your answers.

Please raise your hand if you have questions.

**Part II.** [EXPERIMENTER SAYS IT OUT LOUD]. There are 4 questions in this part and each question consists of several rows. The instructions for the Part II are the same as the instructions for Part I of the experiment. You may proceed and answer the questions in Part II.

**Part III.** [EXPERIMENTER SAYS IT OUT LOUD]. There are 2 questions in this part. As you will see, these questions differ from the ones you have answered before. The instructions will appear on your computer screens. Please read those instructions carefully and answer the questions. You may proceed.

**Risk 1.** You are endowed with 100 points. Each point is worth 10 cents, so you are endowed with \$10. You can choose to invest any amount between 0 and 100 points in a risky project. The remaining amount (points not invested in the risky project) is yours to keep. The risky project has a 50% chance of success:

- If the project is successful, you will receive 2.5 times the amount you chose to invest.
- If the project is unsuccessful, you will lose the amount invested.

Please choose the amount you want to invest in the risky project. Note that you can pick any amount between 0 and 100 points, including 0 or 100.

**Risk 2.** You are endowed with 100 points. Each point is worth 10 cents, so you are endowed with \$10. You can choose to invest any amount between 0 and 100 points in a risky project. The remaining amount (points not invested in the risky project) is yours to keep. The risky project has a 40% chance of success:

- If the project is successful, you will receive 3 times the amount you chose to invest.
- If the project is unsuccessful, you will lose the amount invested.

Please choose the amount you want to invest in the risky project. Note that you can pick any amount between 0 and 100 points, including 0 or 100.

**Part IV.** There are **3 questions** in this part. Each question consists of several rows. In each row, there are two options: the Left Option and the Right Option. Here is an example of a question with 5 rows:

50% chance of \$8 50% chance of \$5	\$5	<input type="text" value="10"/>
50% chance of \$8 50% chance of \$5	\$6	<input type="text"/>
50% chance of \$8 50% chance of \$5	\$6.50	<input type="text"/>
50% chance of \$8 50% chance of \$5	\$7	<input type="text"/>
50% chance of \$8 50% chance of \$5	\$8	<input type="text" value="0"/>

Just like before, the **Left option is always the same** in every row, while the **Right option changes**, getting better and better as you go down the rows.

**In each row, your task is to indicate in the box an (integer) number between 0 and 10.** This number determines the probability with which you get the Left and the Right options:

- If you select 10, you get the Left option **for sure** (probability 100% on Left option)
- If you select 9, you get the Left option with probability 90%, the Right option with probability 10%
- If you select 8, you get the Left option with prob. 80%, the Right option with probability 20%
- If you select 7, you get the Left option with probability 70%, the Right option with probability 30%
- ...
- If you select 2, you get the Left option with probability 20%, the Right option with probability 80%
- If you select 1, you get the Left option with probability 10%, the Right option with probability 90%
- If you select 0, you get the Right option **for sure** (probability 0% on Left option)

In general, the higher the number, **the higher the probability you receive the option on the Left**. Which option you receive will be determined by the computer following the number you specified.

Like in Part I, in some of rows one of the two options pays at least as much as the other. This is the case in the **first** and the **last row** of our example above. In these cases, the numbers 10 or 0 will be pre-entered for you. You can change these numbers if you like.

Recall that there are **no right or wrong answers**, we are only interested in studying **your preferences**; and that there are only **3 questions** in this part, so please think carefully about your answers. Please raise your hand if you have questions.

**Part V.** [EXPERIMENTER SAYS IT OUT LOUD]. This part of the experiment consists of a series of short questions. The instructions for each question are on your computer screens. Please read those instructions carefully and answer the questions. You may proceed.

**IQ 1 - IQ 6.** Subjects were asked to answer 6 IQ questions from the ICAR database (Condon and Revelle, 2014), three Matrix reasoning ones (reminiscent of Raven tests) and three Three-dimensional rotation; these are the same tasks used in (Chapman et al., 2019a). In each of these questions, subjects are presented with the visual geometric design with a missing piece. The task is to find the missing piece.

**Overconfidence 1.** Think about the last 6 puzzles you solved. How many of them do you think you answered correctly?

**Overconfidence 2.** Think again about the last 6 puzzles. Now think about 100 typical people in the United States. Where do you think you rank in terms of how many correct answers you got? For example,

- if you think you got the most correct, you should answer 1;
- if you think you got the least correct, you should answer 100.

**Overconfidence 3.** Think about the wired telephone (landline). What year was the telephone invented? We are interested in your best guess, so please do not look this up if you do not know. Please type the year in which the wired telephone was invented.

How confident are you of your answer to the previous question, in which we asked you to specify the year in which the wired telephone was invented?

- No confidence at all
- Not very confident
- Somewhat unconfident

- Very confident
- Certain

What do you think the probability is (from 0%, or no chance, to 100%, or certainty) that your answer to the question in what year the wired telephone was invented is within 25 years of the correct answer?

Please type the number between 0 and 100 indicating the percentage chance that your answer is within 25 years of the correct answer.

**CRT 1.** A bat and a ball cost \$1.10 in total. The bat costs \$1.00 more than the ball. How much does the ball cost in cents? If you answer correctly, you receive 10 cents. Please enter your answer in cents.

**CRT 2.** If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? If you answer correctly, you receive 10 cents. Please enter your answer in minutes

**CRT 3.** In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? If you answer correctly, you receive 10 cents. Please enter your answer in days